Unravelling the mysteries of the milk protein β-lactoglobulin

---

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Biochemistry

2017

Jennifer Crowther

University of Canterbury

Abstract

The over-arching aim of this thesis is to characterise and compare the biophysical and physicochemical properties of the bovine and caprine orthologues of β-lactoglobulin (βlg). This knowledge is vital to understand the different responses of cow and goat milk to digestion and processing, and differences in the immunogenicity of these milks.

Caprine and bovine βlg were successfully recombinantly expressed using specialised *E. coli* Origami cells conducive to the formation of disulfide bonds. Co-expression with a disulfide bond isomerase encourages the correct disulfide bonds to form. Correctly folded βlg exhibits marked stability at low pH values, thus misfolded βlg and contaminating proteins can be effectively removed by reducing the pH to below 3 and adding moderate concentrations of salt. Circular dichroism spectra and mass spectrometry confirm that each of the βlg proteins are expressed as full-length, folded constructs.

βLg contributes to milk protein allergies in humans, however, goat milk carries a lesser allergenic burden than cow milk. This may be due to differences in the structures of bovine and caprine βlg, or due to differences in the extent of digestion of these proteins. Elucidation of an ultra-high resolution crystal structure of caprine βlg reveals that it shares a very similar structure with that of the bovine variants of βlg. This suggests that antibodies raised against bovine βlg would be capable of cross-reacting with caprine βlg.

Analytical ultracentrifugation studies of the solution structure of caprine βlg indicate that this protein displays a difference in self-association behaviour than bovine βlg at low pH values. Caprine βlg participates in a monomer-dimer self-association at pH values from 3.5 to 7.5, while at pH 2.5 (and 150 mM NaCl) this protein is predominantly monomeric. Bovine βlg, on the other hand, is in a monomer-dimer self-association under these conditions at pH 2.5. This has the potential to affect how these proteins are presented to the immune system during digestion.

Bovine and caprine βlg display a similar level of resistance towards denaturation by urea, as assessed by circular dichroism, while molecular dynamics simulations suggest that the conformational fluctuations of these proteins occur to a similar extent and are localised to the same regions. These findings suggest that these proteins will be recognised by the immune system and acted on by digestive enzymes in the same way.

Bovine βlg is involved in the fouling of surfaces when cow milk is heat-treated. Goat milk is increasing in popularity as an alternative to cow milk, yet far less is known about the thermal denaturation

processes of caprine βlg. An analysis of the thermal denaturation behaviour of caprine βlg, utilising circular dichroism spectroscopy, indicates a dependence of the thermal stability on pH that differs from that of bovine βlg. This is likely a result of the lesser net negative charge of caprine βlg compared to bovine βlg.

Despite being intensively studied for over 50 years, the physiological role of βlg in milk is still a mystery. Sedimentation velocity analyses of fluorescently-labelled βlg proteins provide evidence for a protein-protein interaction between βlg and a component in milk. This is the first time that interactions involving βlg within milk have been investigated, made possible by the use of fluorescence-detection analytical ultracentrifugation. The co-elution of immunoglobulin proteins with βlg from milk following size-exclusion chromatography supports the hypothesis presented that βlg is capable of interacting with immunoglobulins in milk. The physiological role of βlg in milk may be to bind and protect immunoglobulins from digestion as they traverse the digestive tract. This would leave them available for absorption to support the developing immune system of new-born ruminants.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 1: Parts of this chapter have been published in the book chapter: Jennifer M. Crowther, Geoffrey B. Jameson, Alison J. Hodgkinson and Renwick C.J. Dobson (2016). Structure, Oligomerisation and Interactions of β-Lactoglobulin, Milk Proteins - From Structure to Biological Properties and Health Aspects, Dr. Isabel Gigli (Ed.), InTech, DOI: 10.5772/62992.

Please detail the nature and extent (%) of contribution by the candidate:

The candidate wrote this book chapter herself, with proof-reading carried out by her three co-supervisors (the remaining authors).

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: Ren Dobson
Signature:
Date: 5/4/17

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 4: Part of the work presented in this chapter has been published in Crowther, J., Lasse, M., Suzuki, H., Kessans, S., Loo, T., Norris, G., Hodgkinson, A., Jameson, G., Dobson, R. (2014) Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin. FEBS Letters, 588, 3816-3822.

Please detail the nature and extent (%) of contribution by the candidate:

The candidate wrote this journal article herself, and the article is based on experimental work carried out solely by the candidate.

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: Ren Dobson
Signature:
Date: 5/4/17

Acknowledgements

First and foremost my sincere thanks go to my supervisor, Renwick Dobson. Your unwavering support and enthusiasm has undoubtedly seen me through the last four years. Thank you for providing me with wonderful opportunities to travel overseas and attend international conferences, and for supporting me in my future career endeavours. And thanks for all the beers! Thanks also to the other members of my PhD committee, Geoff Jameson and Alison Hodgkinson, your keen insight and sharp eye for detail have been hugely appreciated.

Thanks to Tom Laue for hosting me during my time spent at the University of New Hampshire. I really enjoyed learning some of the finer intricacies of analytical ultracentrifugation. Thank you and Jan*et al*so for so kindly opening your home to me and making me feel welcome while I was in America.

Thank you Jane Allison for taking the time to get me up and running to carry out my own molecular dynamics simulations, and for your ongoing support. I have come to love computing far more than I ever thought I would!

Special thanks goes to my lab comrades: Rach, Kat, Chris, JD, Jenna and Amy, and to all those that have been a part of the Level 6 lab group over the years. You have made it a fun place to come in to work each day and I appreciate all the support you have given me.

Thank you to my incredibly supportive family for all of the love and encouragement you have shown throughout. Antony, you have inspired me to finish and Mum and Dad, you've been with me every step of the way.

Last, but most certainly not least, my heartfelt gratitude goes to my tremendously supportive fiancé Mason. Thank you for your enduring patience, for picking me up when I was feeling low, and for always encouraging me to stick with it. I most certainly could not have done this without you.

Table of contents

Title page..... I
Abstract..... III
Co-Authorship Forms..... V
Acknowledgements..... VII
Table of contents..... IX
Abbreviations..... XIV

1. Introduction 1

1.1 Overview 1
1.2 The composition of cow and goat milk 2
1.3 Bovine βlg is a proven allergen in cow milk 5
1.4 Goat milk is digested more efficiently than cow milk 6
1.5 The response of βlg to heat treatment can affect milk processing 7
1.6 βLg belongs to the lipocalin family 9
1.6.1 Structural features of bovine βlg 10
1.7 The physiological function of βlg remains unknown 12
1.7.1 Function of lipocalins 12
1.7.2 There are several proposed roles of βlg 12
1.8 Research aims of this dissertation 14
1.9 References 16

2. Methods 22

2.1 Experimental reagents 22
2.1.1 Chemical reagents 22
2.1.2 Biological reagents 22
2.2 Microbiology 23
2.2.1 Aseptic technique 23
2.2.2 Media 23
2.2.3 Bacterial strains 23
2.2.4 Antibiotics 24
2.2.5 Preparation of competent cells 24
2.2.6 Transformation of competent cells 25
2.2.7 Bacterial cultures 25
2.2.8 Glycerol stocks 26
2.3 Molecular Biology 26
2.3.1 Plasmid isolation from transformed *E. coli* 26
2.3.2 Restriction enzyme digestion 26
2.3.3 Agarose gel electrophoresis 26
2.3.4 Ligation 27
2.3.5 Sequencing 27
2.3.6 Bioinformatic analyses 28
2.4 Protein Expression 28
2.4.1 Small-scale protein expression 28
2.4.2 Large scale protein expression 29
2.5 Protein Purification 29
2.5.1 Anion exchange chromatography 29
2.5.2 Salt fractionation 30
2.5.3 Size exclusion chromatography 30
2.5.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis 30
2.5.5 Protein quantification 31

2.5.6 Theoretical isoelectric point estimations 32
2.6 Biophysical and Physicochemical Characterisation Methods 32
2.6.1 Mass spectrometry 32
2.6.2 Denaturation experiments 32
2.6.3 Analytical ultracentrifugation 34
2.6.4 X-ray crystallography 36
2.6.5 Small-angle X-ray scattering 37
2.6.6 Molecular dynamics simulations 38
2.7 References 40
3. Cloning, Expression and Purification of Bovine and Caprine β-Lactoglobulin 43
3.1 Introduction 43
3.1.1 Recombinant expression of bovine βlg 43
3.1.2 Disulfide bond formation is required for successful recombinant expression of βlg 44
3.1.3 Overview 44
3.2 Results and discussion 45
3.2.1 Recombinant expression of bovine and caprine βlg 45
3.2.2 Purification of bovine and caprine βlg 49
3.2.3 Mass spectrometry confirms the correct protein products 50
3.2.4 Circular dichroism spectroscopy confirms that each βlg protein is folded 51
3.3 Summary 55
3.4 References 56
4. A Structural Characterisation of Caprine β-lactoglobulin . 59
4.1 Introduction 59
4.1.1 Bovine βlg contributes to cow milk allergies 59
4.1.2 Quaternary structure of bovine βlg 61

4.1.3 Overview 64
4.2 Results and discussion 65
4.2.1 Quaternary state characterisation of caprine βlg by analytical centrifugation 65
4.2.2 Atomic structure of caprine βlg 70
4.2.3 Solution structure of bovine and caprine βlg 77
4.2.4 Comparison of βlg orthologue structures 81
4.3 Summary 88
4.4 References 90
5. Stability and Dynamics of Caprine β-lactoglobulin 94
5.1 Introduction 94
5.1.1 Caprine βlg is digested more efficiently than bovine βlg 94
5.1.2 Cow and goat milk respond differently to heat treatment 95
5.1.3 The different behaviours of bovine and caprine βlg may be due to the stability and dynamics of these proteins 95
5.1.4 Overview 96
5.2 Results and discussion 97
5.2.1 Assessing the conformational stability of βlg proteins 97
5.2.2 Assessing the dynamic nature of βlg proteins 110
5.3 Summary 124
5.4 References 126
6. Interactions between β-lactoglobulin and milk components 130
6.1 Introduction 130
6.1.1 A definitive role for βlg is yet to be ascribed 130
6.1.2 Fluorescence-detection analytical ultracentrifugation 131
6.1.3 Overview 131

6.2 Results and discussion 132
6.2.1 Part 1: Bovine and caprine βlg interact with a component in cow and goat milk 132
6.2.2 Part 2: Identification of the binding partner within milk 142
6.3 Summary 149
6.4 References 150
7. Conclusions and future directions 152
7.1 Overview 152
7.2 Bovine and caprine βlg contribute to milk allergies 152
7.3 βLg is involved in the response of milk to heat treatment 153
7.4 The physiological role of βlg is yet to be defined 154
7.5 Future directions 155
7.6 References 156
8. Appendix 158

Abbreviations

%	Percent
°	Degree
°C	Degrees Celsius
µg	Microgram
µL	Microliter
µm	Micrometre
µM	Micromolar
µs	Microsecond
Å	Angstrom
A_{280}	Absorbance at 280 nm
A_{max}	Absorbance at wavelength of maximum absorbance
AUC	Analytical ultracentrifugation
c(*M*)	Continuous mass distribution
c(*s*)	Continuous sedimentation coefficient distribution
CD	Circular dichroism
CF	Correction factor
cm	Centimetre
CPM	7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin
Da	Daltons
D_{max}	Maximum dimension of the particle
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DsbC	Disulfide bond isomerase C
DSF	Differential scanning fluorimetry

DTT	Dithiothreitol
E. coli	*Escherichia coli*
Fab	Antigen-binding fragment
Fc	Constant fragment
FITC	Fluorescein isothiocyanate
fs	Femtosecond
g	Grams
g	Gravity
HCl	Hydrochloric acid
HT	High tension
I_0	Scatter intensity at the origin
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IPTG	Isopropyl-β-D-thiogalactopyranoside
K	Kelvin
K_D	Dissociation constant
kDa	Kilodaltons
kJ	Kilojoule
L	Litre
LB	Luria-Bertani
M	Molar
m	Metre
MES	2-(N-morpholino)ethanesulfonic acid
mg	Milligram
mL	Millilitre

mM	Millimolar
mm	Millimetre
MRE	Mean residue ellipticity
MRW	Mean residue weight
NaCl	Sodium chloride
ng	Nanogram
nL	Nanolitre
nm	Nanometre
nM	Nanomolar
NPT	Normal pressure and temperature
ns	Nanosecond
OD_{600}	Optical density at 600 nm
$P(r)$	Pair-wise distance distribution
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pIgR	Polymeric immunoglobulin receptor
ps	Picosecond
q	Scattering vector
R_g	Radius of gyration
rmsd	Root mean square deviation
rmsf	Root mean square fluctuations
rpm	Revolutions per minute
s	Sedimentation coefficient
S	Svedberg
$s_{20,w}$	Sedimentation coefficient adjusted to 20° in water
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

sIgA	Secreted IgA
SPC	Simple point charge
TAE	Tris-acetate-ethylene diamine tetra acetic acid
T_m	Melting temperature
UHT	Ultra-high temperature
UV	Ultraviolet
UV/Vis	Ultraviolet/Visible
V	Volts
v/v	volume per volume
VMD	Visual Molecular Dynamics
w/v	weight per volume
βlg	β-lactoglobulin
ε	Extinction coefficient
θ	Ellipticity

Chapter One

1. Introduction

Parts of this chapter have been published in the book chapter: Jennifer M. Crowther, Geoffrey B. Jameson, Alison J. Hodgkinson and Renwick C.J. Dobson (2016). Structure, Oligomerisation and Interactions of β-Lactoglobulin, Milk Proteins - From Structure to Biological Properties and Health Aspects, Dr. Isabel Gigli (Ed.), InTech, DOI: 10.5772/62992.

1.1 Overview

Cow milk is widely available for consumption around the world. However, goat milk is increasing in popularity as an alternative to cow milk due to its perceived health benefits, enhanced digestibility, and lesser allergenic burden (Haenlein, 2004; Jandal, 1996; Lara-Villoslada *et al*., 2004). β-Lactoglobulin (βLg) is a protein found in the whey fraction of both cow and goat milk, yet it is absent in the milk of humans (Kontopidis *et al*., 2014). As a result, βlg is one of the proteins that contribute to cow milk allergies (Kapila *et al*., 2013).

Due to its abundance in cow milk and the relative ease of its purification, bovine βlg has served as a model protein for countless biophysical studies of protein folding, stability and self-association (Sawyer and Kontopidis, 2000). While this has created an extensive literature on the nature of bovine βlg, very little is known about the caprine orthologue of βlg found in goat milk.

Understanding the behaviour of this milk protein is of particular interest to the dairy industry given the potential of βlg to affect the processing and manufacture of milk products; for example, bovine βlg aggregation upon heat treatment is known to contribute to the fouling of

heat exchangers during the processing of milk (de Jong, 1997). It is of value to understand the behaviour of βlg from different species' milks in order to understand how this may influence the processing, digestion, and immunogenicity of these various milks and their products.

Despite being studied for over 50 years, the physiological role of βlg remains a mystery. βLg belongs to the lipocalin family of proteins, most of which have roles that involve ligand binding (Flower, 1996). Its established ability to bind molecules *in vitro* has prompted speculation that βlg is involved in the transport of insoluble and/or chemically sensitive molecules between mother and offspring. However, it is necessary to draw the distinction between demonstrating binding of a molecule *in vitro* and identifying an endogenous ligand that translates to a physiological role of βlg *in vivo*.

This thesis investigates the biophysical and physicochemical properties of caprine (goat) βlg, towards identifying differences between the caprine and bovine orthologues of βlg. This information is essential in order to understand how these proteins may affect the processing, digestion and immunogenicity of these milks. Interaction studies aim to identify interactions occurring between βlg proteins and other components in milk, which may provide insight into the as-yet unknown physiological role of this enigmatic protein.

1.2 The composition of cow and goat milk

Cow milk is composed primarily of casein proteins and whey proteins (O'Donnell *et al*., 2004). The casein proteins include α_{s1}-casein, α_{s2}-casein, β-casein and κ-casein. These constitute around 75% of the proteins in milk. The rest are the whey proteins, β-lg and α-lactalbumin, of which βlg dominates. The presence of numerous genetic variants and various post-translational modifications, as well as the background of low abundance proteins add to the complexity of the milk proteome.

Goat milk contains the same major proteins as cow milk (α-, β- and κ-caseins, α-lactalbumin and βlg) but their genetic polymorphisms and distribution differ (Haenlein, 2004). βLg is still the major whey protein in goat milk (Hinz *et al*., 2012). The largest difference in the protein composition of cow and goat milk is in the casein composition. In cow milk α_{s1}-casein is the

dominant casein. While levels of α_{s1}-casein vary between goat breeds, the levels of this protein are usually lower than in cow milk or are completely devoid, making α_{s2}-casein the major casein of goat milk (Haenlein, 2004). This affects the digestibility and the cheese making properties of goat milk.

An often overlooked feature of goat milk is its lipid content. The lipid composition of goat milk differs significantly from that of cow milk, being much higher in medium-chain triglycerides (Haenlein, 2004). In fact three of these medium-chain triglycerides (C6: caproic acid, C8: caprylic acid and C10: capric acid) have been named after goats due to their abundance in goat milk. Medium-chain triglycerides have become established medical treatments for a range of digestive and malnutrition disorders due to their metabolic advantage of providing immediate energy as opposed to being deposited in adipose tissue (López-Aliaga *et al*., 2010). Goat milk is also higher in monounsaturated and polyunsaturated fatty acids which, along with medium-chain triglycerides, are known to be beneficial for human health (Haenlein, 2004).

At least eleven variants of bovine βlg have been found in cow milk, with A and B the most common (Ballester *et al*., 2005). Variant B differs from A by two amino acid substitutions: D64G and V118A (Figure 1.1). Only one protein variant of caprine βlg has been described in goat milk (Ballester *et al*., 2005). Caprine βlg contains the same substitutions as bovine βlg B relative to bovine βlg A: D64G and V118A. Caprine βlg also differs from both variants of bovine βlg at a further six sites: L1I, D53N, D130K, S150A, E158G and I162V (Figure 1.1) (Folch *et al*., 1994). βLg is also found in the milk of most other mammals yet is absent in humans, camels, rodents and lagomorphs (e.g. rabbits) (Kontopidis *et al*., 2004). While the level of sequence similarity decreases as species become more divergent it is evident that there are several, potentially functionally significant, residues which are highly conserved (Figure 1.1).

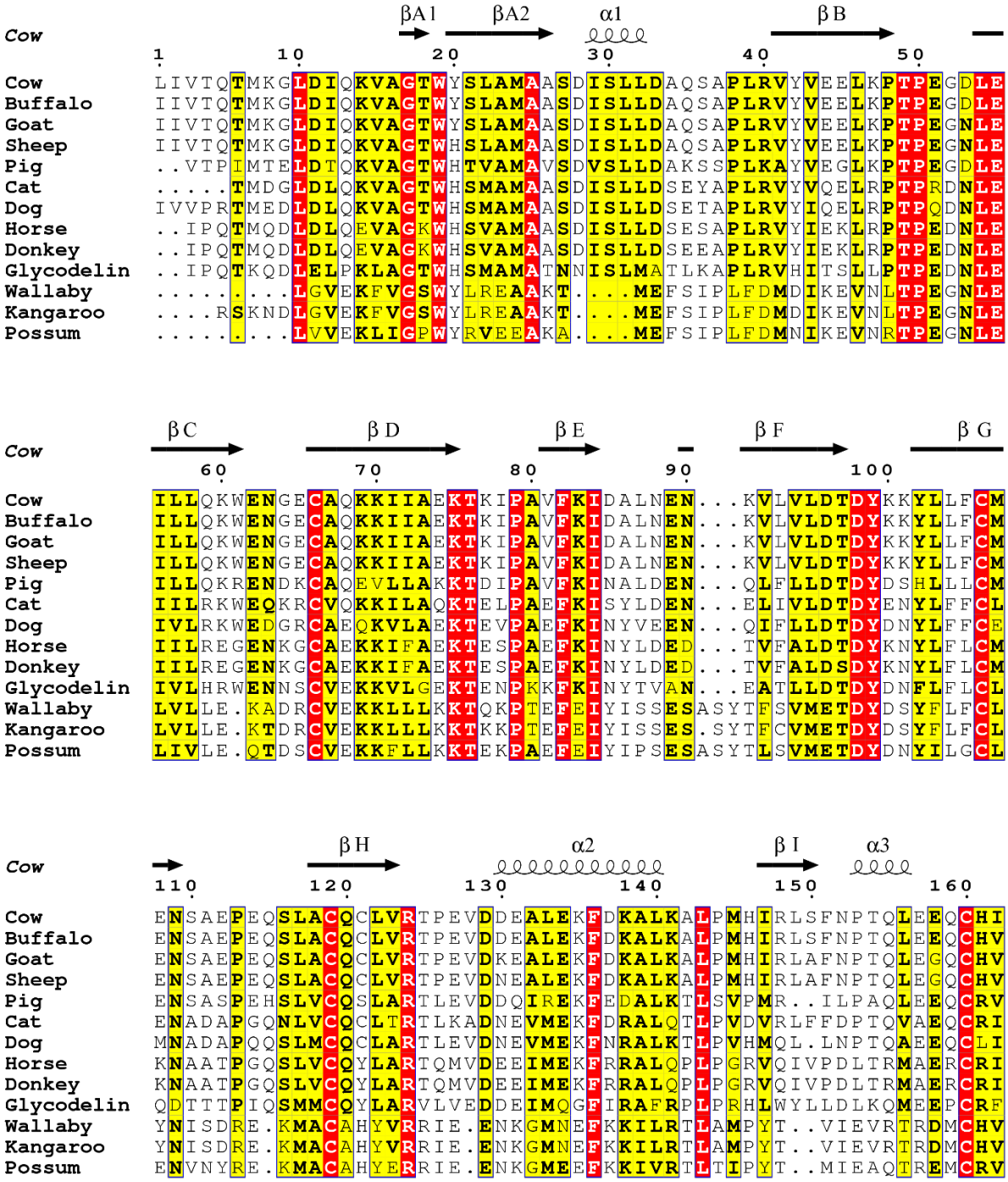


Figure 1.1: Sequence alignment of βlg from various species. The numbering and secondary structure is as according to bovine βlg. The highlighted yellow residues indicate areas of moderate conservation, while red indicates a high level of conservation. The UniProt accession numbers for each sequence are as follows: Cow (A variant) P02754, Buffalo P02755, Goat P02756, Sheep P67976, Pig P04119, Cat P33688, Dog P33685, Horse P02758, Donkey P13613, Glycodelin P09466 (closest homologue in humans to βlg), Wallaby Q29614, Kangaroo P11944 and Possum Q29146. Sequences are aligned using Clustal Omega (Sievers and Higgins, 2013) and figure is generated using ESPript (Robert and Gouet, 2014).

1.3 Bovine βlg is a proven allergen in cow milk

Understanding the behaviour and characteristics of milk components is crucial for understanding the underlying causes of milk allergies. Milk allergy is one of the most common food allergies, with a prevalence of 2-3% in infants. Most children will outgrow their allergy within the first years, although they retain a greater risk to develop respiratory allergies such as hay-fever or asthma (Saarinen *et al*., 2005). βLg is not present in the milk of humans and has been identified as one of the main immunogenic proteins contributing to cow milk allergies (Haenlein, 2004).

While some studies have indicated that patients with cow milk allergies show improvement of symptoms when they consume goat milk (Haenlein, 2004), others have suggested that children with cow milk protein allergies are also sensitised to goat milk proteins, suggesting cross reactivity between the proteins (Lara-Villoslada *et al*., 2004). The discrepancy between findings may have been due to differences in the underlying causes of milk allergy. In children clinically diagnosed with Immunoglobulin E (IgE)-mediated cow milk protein allergy, both cow milk and goat milk were able to elicit positive skin test responses and allergic reactions (urticaria, rhinitis, vomiting etc.) when administered orally (Bellioni-Businco *et al*., 1999). However, a greater amount of goat milk (five times as much) was required to elicit the same response as cow milk.

While goat milk may not be a suitable alternative for infants already showing signs of cow milk allergy, goat milk may be a less immunogenic option to be introduced into the diet of non-sensitised children. Lara-Villoslada *et al*. (2004) carried out the first animal study comparing cow milk and goat milk as the first protein sources after breast-feeding. They used a mouse model that exhibits the characteristics of type 1 hypersensitivity reactions (i.e. an allergic immune response involving IgE antibodies as opposed to the IgA, IgG or IgM antibodies involved in normal infectious immune responses) in order to compare the allergenicity of cow milk and goat milk. Their results suggest that goat milk is less immunogenic than cow milk, as it induces a weaker allergic response.

The fact that goat milk carries a lesser allergenic burden may be due to structural differences between the protein allergens within each milk, due to species-specific sequence variations. These changes may alter the surface structure, the part of the molecule that interacts with

antibodies, and reduce antibody affinity (Aalberse, 2006). The proteins may still share enough structural similarity to be cross-reactive with one another. Alternatively, the amount of allergen being available to pass through the walls of the digestive tract in its native, allergenic state may be reduced. This could occur as a result of the enhanced digestibility of goat milk (see section 1.4) allowing more complete digestion of βlg and other allergens. More in-depth knowledge of the structure and stability of caprine βlg may provide a deeper understanding of this.

1.4 Goat milk is digested more efficiently than cow milk

Anecdotal reports of the medical and nutritional benefits of goat milk abound, yet few scientific studies have been published in peer-reviewed journals. Those that have suggest that goat milk is more easily, and thoroughly, digested than cow milk (Almaas *et al*., 2006; Ceballos *et al*., 2009). The difference in digestibility between cow and goat milk is largely due to the differing fat composition of each milk (López-Aliaga *et al*., 2010). The fat globules of goat milk are smaller than those in cow milk and do not naturally aggregate upon cooling. This is due to the lack of the protein agglutinin, which is responsible for the aggregation of fat globules in cow milk (Park *et al*., 2007). The natural homogenisation of goat milk makes it easier to digest as it allows better access by lipases in the gut.

Protein composition further explains the enhanced digestibility of goat milk. The levels of α_{s1}-casein in goat milk vary between breeds, yet are commonly much lower than in cow milk (Ceballos *et al*., 2009). The lower levels of this casein protein result in a softer curd formation in the stomach, which facilitates the access of proteolytic enzymes within the digestive tract.

Simulated *in vitro* digestion of cow and goat milk using human proteolytic enzymes has shown that caprine milk proteins are digested faster, and more completely, than bovine milk proteins (Almaas *et al*., 2006). This is particularly evident for βlg. After treatment with human gastric juice and duodenal juice only 23% of βlg remained undigested in goat milk, while 83% remained in cow milk. It has been suggested that caprine βlg exhibits a reduced conformational stability compared to the bovine A and B variants of βlg (Alexander and Pace, 1971) and is therefore less resistant to denaturation by treatment with urea. Whether the differential digestibility

seen between these proteins is due to protein stability differences or due to compositional differences between the milks remains to be seen.

1.5 The response of βlg to heat treatment can affect milk processing

Heat treatment of milk and dairy products is often necessary to eliminate pathogenic organisms, but it can also be used to enhance desirable properties of dairy products, such as taste and texture (Anema and McKenna, 1996; Pesic *et al*., 2016). However, the fouling of surfaces during heat treatment of milk is a costly problem (de Jong, 1997). Proteins and minerals form heat-induced deposits on the walls of equipment which lead to undesirable side effects, such as decreased heat transfer, product loss and increased cleaning costs.

It is necessary to understand the reactions occurring during the heat treatment of milk in order to optimise process conditions and to attain desired functional properties in milk products. Milk is a complex mixture, and thus it is difficult to define the exact mechanisms and reactions occurring between milk components during heat processing. However, the aggregation of βlg during heat treatment has been shown to play a dominant role in the fouling process (de Jong, 1997). Therefore, considerable effort has been dedicated to understanding the thermal denaturation behaviour of βlg.

The process of thermal denaturation and aggregation of bovine βlg has been thoroughly investigated (Anema and McKenna, 1996; Galani and Apenten, 1999; Roefs and Kruif, 1994). βLg exists in milk as a dimer of two, non-covalently linked, subunits. Within each subunit there are two disulfide bonds and one free thiol group, which is buried in the native structure. The first step of thermal denaturation is dissociation into monomers (Roefs and Kruif, 1994). Following this the monomers begin to unfold. This unfolding reveals the buried thiol group which is then able to react via disulfide exchange with one of the two intramolecular disulfide bonds of another βlg monomer. This forms an intermolecular disulfide bond and a new reactive thiol. This step can be repeated many times forming large aggregates.

Reactive βlg can also react with other milk components that contain disulfide bonds, such as α-lactalbumin (Hong and Creamer, 2002), bovine serum albumin (Havea *et al*., 2001) and κ-

casein (Anema and Stanley, 1998; Henry *et al*., 2002), or with components on the heat exchanger surfaces leading to protein deposition and fouling of these surfaces (de Jong, 1997). Complexes between denatured whey proteins and κ-casein may exist either as micelle-associated complexes or as soluble complexes in the serum phase of milk (Pesic *et al*., 2016). Different factors such as pH, temperature, heating time and the composition of milk can affect the formation and distribution of these complexes (Anema and McKenna, 1996). This in turn can affect the functional properties of dairy products such as the taste and texture of cheese and yoghurt, the gelation of UHT (ultra-high temperature) milk upon storage, and the properties of skim milk powders (Pesic *et al*., 2016).

Far fewer studies have investigated the response of goat milk to heat treatment. Those that have suggest that goat milk displays a heat stability dependence on pH that differs from that of cow milk (Anema and Stanley, 1998; Bouhallab *et al*., 2002; Montilla and Calvo, 1997; Zadow *et al*., 1983). The heat coagulation time versus pH profile at 140°C indicates that cow milk exhibits a heat stability maximum at pH 6.7, whereas goat milk has a maximum nearer to pH 6.9 (Anema and Stanley, 1998). Therefore, given that the pH of cow milk is around 6.7 and that of goat milk is around 6.6, goat milk is less heat stable than cow milk at its natural pH. These findings agree with those of Zadow *et al*. (1983); if goat milk undergoes UHT processing (140°C for 3 seconds) below pH 6.9, severe sedimentation of casein and whey proteins occurs. This can be remedied by adjusting the pH to above 7. In contrast, the pH of cow milk only needs to be above 6.6 in order to avoid this extensive sedimentation, with rapid sedimentation occurring at pH 6.4. Adding ionic calcium to milk results in the onset of sedimentation at higher pH values, while addition of phosphates increases the heat stability of both cow and goat milk (Montilla and Calvo, 1997; Wang *et al*., 2015; Zadow *et al*., 1983). The increased sensitivity of goat milk to UHT processing is thought to be due, in part, to the higher levels of ionic calcium found in goat milk (Zadow *et al*., 1983).

The work of Henry *et al*. (2002) provided evidence that, as in cow milk, caprine βlg is capable of forming heat-induced disulfide bonds with κ-casein, forming covalently linked complexes between denatured whey proteins and casein micelles. As the pH at which the milk is heat-treated increases, increasing levels of these denatured whey protein – κ-casein complexes dissociate from micelles and form soluble complexes. The increased dissociation destabilises

the micelles and leads to their precipitation (Bouhallab *et al*., 2002). It appears that in cow milk the pH at which whey protein – κ-casein complexes increasingly dissociate from the micelles is pH 6.6, while in goat milk this increase occurs closer to pH 6.8 (Anema and Stanley, 1998; Pesic *et al*., 2012).

These behaviours are likely due to differences in the interactions of denatured whey proteins, including βlg, with κ-casein due to variations in the primary structures of these proteins, the micellar structure, casein composition and the ionic concentration of calcium (Anema and Stanley, 1998; Henry *et al*., 2002; Oldfield *et al*., 2000; Pesic *et al*., 2012). While further research is needed to understand the surface structure of bovine and caprine casein micelles, it also remains to be seen whether the physicochemical characteristics of bovine and caprine whey proteins, such as βlg, influence the different behaviours of these milks upon heat treatment. This knowledge is required in order to optimise milk processing conditions.

1.6 βLg belongs to the lipocalin family

The first reported atomic-resolution structure of βlg, solved by X-ray crystallography for bovine βlg (Papiz *et al*., 1986), showed remarkable similarity to retinol binding protein. βLg was thus added to the family of proteins known as lipocalins. Lipocalin proteins all share a characteristic fold despite great diversity at the sequence level. The conserved lipocalin fold includes an eight-stranded anti-parallel β-sheet that is folded back upon itself to enclose an internal cavity, often termed a calyx (Figure 1.2) (Flower, 1996). This fold makes these proteins well suited to binding a range of ligands with the amino acids lining the calyx determining binding specificity. While they were once simply classified as transport proteins, lipocalins are now known to exhibit vast functional diversity (see section 1.7.1), yet most involve some form of ligand binding (Sawyer and Kontopidis, 2000).

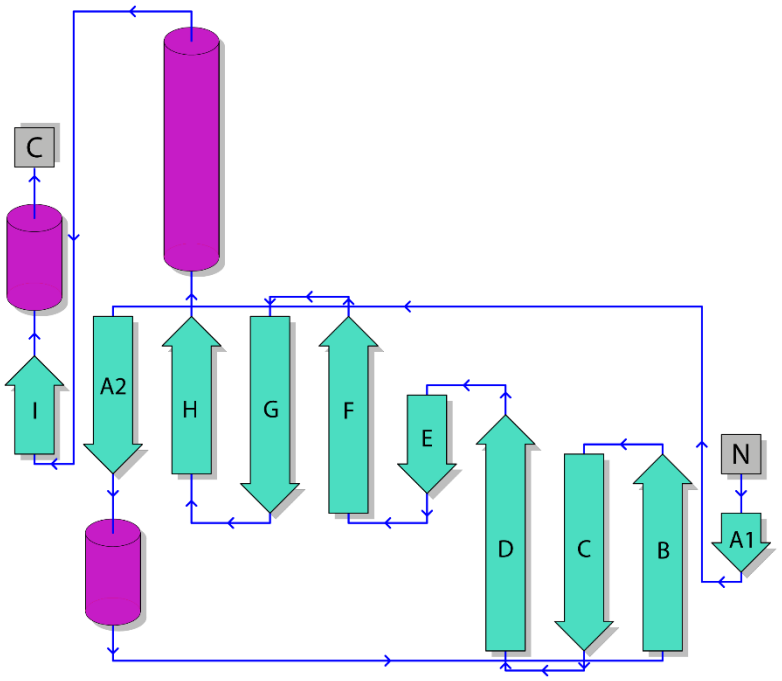


Figure 1.2: Topology diagram showing characteristic features of lipocalin proteins. The eight β-strands A-H form the antiparallel β-barrel. Strand A is kinked (giving rise to strands A1 and A2) and connects the β-sheet comprising strands A1, B, C and D to the β-sheet comprising strands E, F, G, H and A2. An α-helix follows strand H, followed by a ninth β-strand, I, that packs against strand A2. Figure generated by PDBsum (Laskowski *et al.*, 1997) using the structure of bovine βlg A, PDB ID: 1BSO (Qin *et al.*, 1998). (Crowther *et al.*, 2016) © under CC BY 3.0 license.

1.6.1 Structural features of bovine βlg

Numerous high resolution atomic structures now exist for bovine βlg. As shown in the three-dimensional cartoon form in Figure 1.3, each subunit of the dimeric protein consists of an antiparallel β-sheet formed by eight β-strands, A to H, wrapped around to form a flattened calyx (Qin *et al.*, 1998). Following the barrel is a three-turn α-helix and a ninth β-strand which, at least for bovine βlg, forms part of the dimer interface. At the commencement of this study there were no reported structures of the caprine orthologue of βlg.

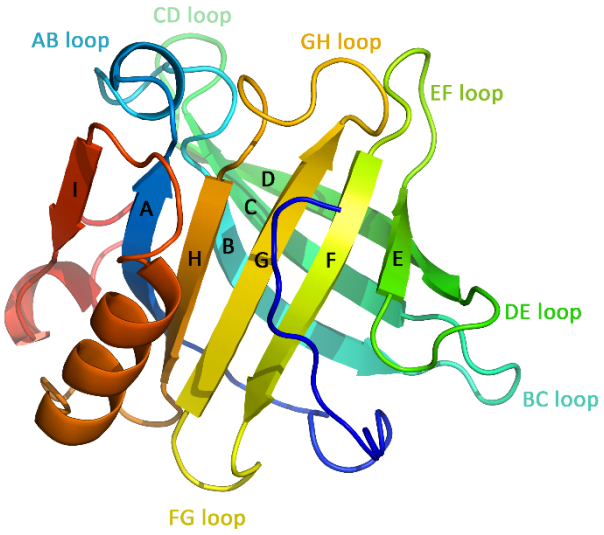


Figure 1.3: Crystal structure of bovine βlg A (PDB ID: 1BSO). Eight β-strands (A-H) form the central antiparallel β-sheet calyx. The AB, CD, EF and GH loops are located at the open end of the calyx. The calyx is flanked on its outer surface by a three-turn α-helix and a ninth β-strand, I. The polypeptide chain is shown coloured in a rainbow, beginning in blue at the N-terminus and ending in red at the C-terminus. Figure produced using the PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC. (Crowther *et al.*, 2016)

Bovine βlg exists in a monomer-dimer equilibrium. While bovine βlg dissociates into monomers at low pH (< pH 3 and low salt concentrations) (Mercadante *et al.*, 2011), the secondary and tertiary structure are maintained (Kuwata *et al.*, 1999; Uhrínová *et al.*, 2000). Contributing to the stability of the monomer is the presence of two disulfide bonds, Cys106-Cys119 and Cys66-Cys160 (Ponniah *et al.*, 2010). The first links β-strands G and H, while the second tethers the C-terminus to the CD loop. These disulfide bonds are required for the correct folding of βlg proteins, a factor which needs to be considered when producing these proteins recombinantly (Ponniah *et al.*, 2010) (see Chapter 3). The structural stability at low pH contributes to the resistance of βlg to proteolytic enzymes within the human digestive tract (Maier *et al.*, 2006; Reddy *et al.*, 1988).

1.7 The physiological function of βlg remains unknown

1.7.1 Function of lipocalins

The lipocalin family comprises a wide range of small extracellular proteins. Despite diversity at the sequence level, these proteins share a conserved protein fold; an eight stranded β-barrel enclosing a large cup-shaped calyx. This fold makes lipocalins well suited to binding a range of hydrophobic molecules. While once simply classified as transport proteins, it is now known that lipocalins exhibit great functional diversity. These functions include retinol and pheromone transport (retinol-binding protein and mouse urinary protein), colouration (insecticyanin), olfaction (odorant binding proteins), prostaglandin synthesis (prostaglandin D synthase), cell regulation (quiescence-specific protein) and modulation of immune responses (pregnancy protein 14) (Flower, 1996).

1.7.2 There are several proposed roles of βlg

Due to its similarity to retinol-binding protein, it was originally hypothesised that the biological function of bovine βlg may be to transport vitamin A between mother and calf. This hypothesis was supported by fluorescence spectroscopy binding data (Pérez and Calvo, 1995) and by the fact that the retinol molecule could be modelled into the crystal structure of bovine βlg in a similar position as seen bound in retinol binding protein (Papiz *et al*., 1986). However, since then bovine βlg has been shown to be capable of binding a broad range of small hydrophobic molecules, as demonstrated in various ligand-bound crystal structures. These bound ligands include vitamin D (Kontopidis *et al*., 2004; Yang *et al*., 2008), vitamin A (Kontopidis *et al*., 2002), cholesterol (Kontopidis *et al*., 2004), a range of fatty acids (Wu *et al*., 1999; Loch *et al*., 2013b, 2013a, 2012, 2011) and the fatty acid derivative 12-bromododecanoic acid (Qin *et al*., 1998), as well as more diverse molecules including sodium-dodecyl-sulfate (SDS) (Gutiérrez-Magdaleno *et al*., 2013) and various anaesthetic drugs (Loch *et al*., 2015).

Fatty acids have been seen bound to bovine βlg isolated from cow milk under non-denaturing conditions, but are present in quantities reflecting the fatty acid composition of milk (Pérez *et al*., 1993). The apparent lack of selectivity, makes it less likely that βlg is a specific fatty acid or

vitamin transporter. However, βlg may still be involved in fat metabolism; it is possible that βlg may stimulate a pre-gastric lipase, potentially by binding the fatty acid products and thereby reducing their inhibitory effect on the enzyme (Pérez *et al*., 1992).

The closest human homologue to βlg is pregnancy protein 14 (PP14, also known as glycodelin) (Seppälä *et al*., 2002). Unlike βlg, PP14 is a glycoprotein. This glycosylation is crucial for the various functions of PP14, which include an immunosuppressive activity in the uterus to protect products of the reproductive organs from the immune system. While tempting to suggest an immuno-regulatory role for βlg, there is no evidence to support this, and the lack of glycosylation on βlg makes this unlikely. It is possible that βlg may have diverged from PP14 following a gene duplication event and has since lost any PP14-related activity. This protein may now simply act as a source of amino acids for the offspring of the animals that produce it. However, the resistance of this protein to digestion (Reddy *et al*., 1988) and the high level of sequence conservation seen between βlg orthologues (Figure 1.1) argue against a simple nutritive function.

It has been suggested that intact βlg protein, along with digestion-derived peptides, may possess antimicrobial activity. The intact protein appears to be capable of inhibiting the growth of *Staphylococcus aureus, Streptococcus uberis*, and *Escherichia coli*, bacteria largely responsible for the prevalent and costly disease, bovine mastitis (Chaneton *et al*., 2011). βLg also appears to augment the antimicrobial activity of lactoferrin, a protein with a known role in the defence against mammary gland infections (Hernández-Ledesma *et al*., 2008) . A separate study has described the antimicrobial activity of four peptides derived from βlg following trypsin digestion (Pellegrini *et al*., 2001) . It is possible, therefore, that βlg may have a protective physiological role in new-born calves and/or in the secretory tissue of the mother. Further studies are needed to fully understand the relevance of these findings, and to address the mechanisms underlying these observations of antimicrobial activity.

These findings serve to show that there is still much confusion and uncertainty surrounding the role of βlg within milk. There is a need for the investigation of interactions that occur between βlg and other protein components within milk under physiological conditions, in order to gain insight into the intended purpose of this protein.

1.8 Research aims of this dissertation

Given the interest in goat milk as a potential substitute for cow milk, particularly as a first protein source in infants, it is crucial to have an understanding of the components that constitute each milk. An in-depth knowledge of the structure, solution behaviour and stability of caprine βlg is required to understand how this protein may be presented to the immune system during digestion, and how it may behave during goat milk processing.

In order to obtain pure protein amenable to biophysical and physicochemical characterisation, I have recombinantly expressed and purified the A and B variants of bovine βlg and caprine βlg. Successful recombinant expression of these proteins requires the use of specialised *E. coli* Origami (DE3) cells that accommodate the formation of disulfide bonds. Co-expression with a disulfide bond isomerase ensures that the correct disulfide bonds form. The far-UV and near-UV circular dichroism spectra indicate that caprine βlg adopts a fold similar to that of the bovine A and B variants of βlg. Mass spectrometry confirms that the recombinantly expressed proteins are expressed as full-length constructs.

Given the amino acid differences between the bovine A and B variants of βlg and caprine βlg (as detailed in section 1.2) it was deemed necessary to gain an understanding of the structure of caprine βlg in order to identify any structural changes between orthologues of βlg. To this end, I have elucidated an ultra-high resolution three-dimensional model by X-ray crystallography of caprine βlg. This reveals that caprine βlg shares significant structural similarity with that of the bovine orthologues of βlg. An analysis of the solution structure of βlg proteins as a function of pH has revealed interesting differences between caprine and bovine βlg. In particular, caprine βlg is predominantly monomeric at low pH (pH 2.5), while bovine βlg participates in a monomer-dimer self-association under the same conditions. This has the potential to affect how these proteins are presented to the immune system during the low pH conditions potentially encountered during digestion.

It is important to have an understanding of the relative stabilities of the bovine and caprine orthologues of βlg as this may inform the different responses seen for cow and goat milk towards heat processing and digestion. Here, the conformational and thermodynamic stability of caprine βlg has been compared to that of bovine βlg. While the bovine and caprine

orthologues of βlg exhibit a similar level of resistance toward denaturation by urea, they display a different dependence on pH in terms of their thermal stability. At pH values of less than 6.9 samples of caprine βlg precipitate upon heating. This observation was not made for bovine βlg at any of the pH values examined. This is likely due to the reduced net negative charge of caprine βlg which leads to a greater extent of hydrophobically-driven associations between heat-induced aggregates.

Differences in protein dynamics due to the amino acid sequence substitutions between bovine and caprine βlg may further explain the different responses of these proteins in terms of digestion, processing and immunogenicity. The dynamic nature of these proteins were compared utilising molecular dynamics simulations. The simulations suggest that the monomers of each protein exhibit similar conformational fluctuations in the same regions of each protein, suggesting that the dynamics of each orthologue are not significantly altered. However, they do reveal a high degree of flexibility at the dimer interface and indicate the presence of numerous interactions (hydrogen bonds and salt-bridges) that are not seen in crystal structures. This gives a greater insight into the residues that participate in dimer stabilisation of βlg proteins.

In order to shed light on the elusive physiological role of βlg, there is a need to observe the interactions that occur with this protein in its physiological environment. The interactions that occur between βlg and other milk components has been examined utilising fluorescence-detection analytical ultracentrifugation. The sedimentation of bovine and caprine βlg are both significantly altered in the presence of other components of cow and goat milk, indicating the formation of a complex with another, larger, component within milk. We have hypothesised that this component may be an immunoglobulin. It may be that βlg proteins are capable of protecting these immunoglobulins as they travel through the digestive tract, increasing their availability to ruminant infants. This interaction involving βlg has not been previously reported and the proposed physiological role of this protein has thus far not been proposed. This role would account for the high levels of conservation seen between orthologues of βlg and would explain the abundance of this protein in cow and goat milk, while it is not required in the milk of humans.

1.9 References

Aalberse, R.C., 2006. Molecular mechanisms in allergy and clinical immunology. The Journal of Allergy and Clinical Immunology 106, 228–38.

Alexander, S.S., Pace, C.N., 1971. A comparison of the denaturation of bovine β-lactoglobulins A and B and goat β-lactoglobulin. Biochemistry 10, 2738–2743.

Almaas, H., Cases, A., Devold, T.G., Holm, H., Langsrud, T., Aabakken, L., Aadnoey, T., Vegarud, G.E., 2006. *In vitro* digestion of bovine and caprine milk by human gastric and duodenal enzymes. International Dairy Journal 16, 961–968.

Anema, S., McKenna, A., 1996. Reaction kinetics of thermal denaturation of whey proteins in heated reconstituted whole milk. Journal of Agricultural and Food Chemistry 44, 422–428.

Anema, S., Stanley, D., 1998. Heat-induced, pH-dependent behaviour of protein in caprine milk. International dairy journal 8, 917–923.

Ballester, M., Sánchez, A., Folch, J., 2005. Polymorphisms in the goat β-lactoglobulin gene. Journal of Dairy Research 72, 379–384.

Bellioni-Businco, B., Paganelli, R., Lucenti, P., 1999. Allergenicity of goat's milk in children with cow's milk allergy. Journal of Allergy and Clinical Immunology 1191–1194.

Bouhallab, S., Leconte, N., Graet, L.Y., Garem, A., 2002. Heat-induced coagulation of goat milk: modification of the environment of the casein micelles by membrane processes. Le Lait 82, 673–681.

Ceballos, L., Morales, E., Adarve, G., Castro, J., Martínez, L., Sampelayo, M., 2009a. Composition of goat and cow milk produced under similar conditions and analyzed by identical methodology. Journal of Food Composition and Analysis 22, 322–329.

Ceballos, L., Morales, E., Martínez, L., Extremera, F., Sampelayo, M., 2009b. Utilization of nitrogen and energy from diets containing protein and fat derived from either goat milk or cow milk. Journal of Dairy Research 76, 497–504.

Chaneton, L., Sáez, J., Bussmann, L., 2011. Antimicrobial activity of bovine β-lactoglobulin against mastitis-causing bacteria. Journal of Dairy Science 94, 138–145.

Crowther, J., Jameson, G., Hodgkinson, A., Dobson, R., 2016. Structure, Oligomerisation and Interactions of β-Lactoglobulin. InTech.

De Jong, P., 1997. Impact and control of fouling in milk processing. Trends in Food Science & Technology 8, 401–405.

Flower, D.R., 1996. The lipocalin protein family: structure and function. The Biochemical Journal 318, 1–14.

Folch, J.M., Coll, A., Sanchez, A., 1994. Complete sequence of the caprine beta-lactoglobulin gene. Journal of dairy science 77, 3493–7.

Galani, D., Apenten, R., 1999. Heat-induced denaturation and aggregation of β-Lactoglobulin: kinetics of formation of hydrophobic and disulphide-linked aggregates. International Journal of Food Science & Technology 34, 467–476.

Gutiérrez-Magdaleno, G., Bello, M., Portillo-Téllez, C.M., Rodríguez-Romero, A., García-Hernández, E., 2013. Ligand binding and self-association cooperativity of β-lactoglobulin. Journal of Molecular Recognition 26, 67–75.

Haenlein, G., 2004. Goat milk in human nutrition. Small Ruminant Research 51, 155–163.

Havea, P., Singh, H., Creamer, L.K., 2001. Characterization of heat-induced aggregates of beta-lactoglobulin, alpha-lactalbumin and bovine serum albumin in a whey protein concentrate environment. Journal of dairy research 68, 483–97.

Henry, G., Mollé, D., Morgan, F., Fauquant, J., Bouhallab, S., 2002. Heat-induced covalent complex between casein micelles and beta-lactoglobulin from goat's milk: identification of an involved disulfide bond. Journal of Agricultural and Food Chemistry 50, 185–91.

Hernández-Ledesma, B., Recio, I., Amigo, L., 2008. β-Lactoglobulin as source of bioactive peptides. Amino Acids 35, 257–265.

Hinz, K., O'Connor, P., Huppertz, T., Ross, R., Kelly, A., 2012. Comparison of the principal proteins in bovine, caprine, buffalo, equine and camel milk. The Journal of Dairy Research 79, 185–91.

Hong, Y., Creamer, L., 2002. Changed protein structures of bovine β-lactoglobulin B and α-lactalbumin as a consequence of heat treatment. International Dairy Journal 12, 345–359.

Jandal, J., 1996. Comparative aspects of goat and sheep milk. Small Ruminant Research 22, 177–185.

Kapila, R., Kavadi, P., Kapila, S., 2013. Comparative evaluation of allergic sensitization to milk proteins of cow, buffalo and goat. Small Ruminant Research 112, 191–198.

Kontopidis, G, Holt, C, Sawyer, L, 2004. Invited review: β-lactoglobulin: binding properties, structure, and function. Journal of Dairy Science 87, 785–796.

Kontopidis, G., Gilliver, A., Sawyer, L., 2014. Ovine β-lactoglobulin at atomic resolution. Acta Crystallographica Section F: Structural Biology Communications 70, 1498–1503.

Kontopidis, G., Holt, C., Sawyer, L., 2002. The Ligand-binding Site of Bovine β-Lactoglobulin: Evidence for a Function? Journal of Molecular Biology 318, 1043–1055.

Kuwata, K., Era, S., Hoshino, M., Forge, V., Goto, Y., Batt, C., 1999. Solution structure and dynamics of bovine β-lactoglobulin A. Protein Science 8, 2541–2545.

Lara-Villoslada, F., Olivares, M., Jiménez, J., Boza, J., Xaus, J., 2004. Goat milk is less immunogenic than cow milk in a murine model of atopy. Journal of Pediatric Gastroenterology and Nutrition 39, 354–360.

Laskowski, R.A., Hutchinson, E.G., Michie, A.D., Wallace, A.C., Jones, M.L., Thornton, J.M., 1997. PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends in Biochemical Sciences 22, 488–490.

Loch, J., Bonarek, P., Polit, A., Jabłoński, M., Czub, M., Ye, X., Lewiński, K., 2015. β-Lactoglobulin interactions with local anaesthetic drugs – Crystallographic and calorimetric studies. International Journal of Biological Macromolecules 80, 87–94.

Loch, J., Bonarek, P., Polit, A., Riès, D., Dziedzicka-Wasylewska, M., Lewiński, K., 2013a. Binding of 18-carbon unsaturated fatty acids to bovine β-lactoglobulin—Structural and thermodynamic studies. International Journal of Biological Macromolecules 57, 226–231.

Loch, J., Bonarek, P., Polit, A., Świątek, S., Dziedzicka-Wasylewska, M., Lewiński, K., 2013b. The differences in binding 12-carbon aliphatic ligands by bovine β-lactoglobulin isoform A and B studied by isothermal titration calorimetry and X-ray crystallography. Journal of Molecular Recognition 26, 357–367.

Loch, J., Polit, A., Bonarek, P., Olszewska, D., Kurpiewska, K., Dziedzicka-Wasylewska, M., Lewiński, K., 2012. Structural and thermodynamic studies of binding saturated fatty acids to bovine β-lactoglobulin. International Journal of Biological Macromolecules 50, 1095–1102.

Loch, J., Polit, A., Górecki, A., Bonarek, P., Kurpiewska, K., Dziedzicka-Wasylewska, M., Lewiński, K., 2011. Two modes of fatty acid binding to bovine β-lactoglobulin—crystallographic and spectroscopic studies. Journal of Molecular Recognition 24, 341–349.

López-Aliaga, I., Díaz-Castro, J., Alférez, M.M., Barrionuevo, M., Campos, M.S., 2010. A review of the nutritional and health aspects of goat milk in cases of intestinal resection. Dairy Science & Technology 90, 611–622.

Maier, I., Okun, V., Pittner, F., Lindner, W., 2006. Changes in peptic digestibility of bovine β-lactoglobulin as a result of food processing studied by capillary electrophoresis and immunochemical methods. Journal of Chromatography B 841, 160–167.

Mercadante, D., Melton, L., Norris, G., Loo, T., Williams, M., Dobson, R., Jameson, G., 2011. Bovine β-Lactoglobulin Is Dimeric Under Imitative Physiological Conditions: Dissociation Equilibrium and Rate Constants over the pH Range of 2.5–7.5. Biophysical Journal 103, 303–312.

Montilla, A., Calvo, M., 1997. Goat's milk stability during heat treatment: effect of pH and phosphates. Journal of Agricultural and Food Chemistry 45, 931–934.

Oldfield, D.J., Singh, H., Taylor, M.W., Pearce, K.N., 2000. Heat-induced interactions of β-lactoglobulin and α-lactalbumin with the casein micelle in pH-adjusted skim milk. International Dairy Journal 10, 509–518.

O'Donnell, R., Holland, J.W., Deeth, H.C., Alewood, P., 2004. Milk proteomics. International Dairy Journal 14, 1013–1023.

Papiz, M.Z., Sawyer, L., Eliopoulos, E.E., North, A.C., Findlay, J.B., Sivaprasadarao, R., Jones, T.A., Newcomer, M.E., Kraulis, P.J., 1986. The structure of beta-lactoglobulin and its similarity to plasma retinol-binding protein. Letters To Nature 324, 383–385.

Park, Y., Juárez, M., Ramos, M., Haenlein, G., 2007. Physico-chemical characteristics of goat and sheep milk. Small Ruminant Research 68, 88–113.

Pellegrini, A., Dettling, C., Thomas, U., Hunziker, P., 2001. Isolation and characterization of four bactericidal domains in the bovine beta-lactoglobulin. Biochimica et biophysica acta 1526, 131–40.

Pesic, M., Barac, M., Stanojevic, S., 2016. Heat-Induced Casein–Whey Protein Interactions in Caprine Milk: Whether Are Similar to Bovine Milk? Springer.

Pesic, M.B., Barac, M.B., Stanojevic, S.P., Ristic, N.M., Macej, O.D., Vrvic, M.M., 2012. Heat induced casein–whey protein interactions at natural pH of milk: A comparison between caprine and bovine milk. Small Ruminant Research 108, 77–86.

Ponniah, K., Loo, T., Edwards, P., Pascal, S., Jameson, G., Norris, G., 2010. The production of soluble and correctly folded recombinant bovine β-lactoglobulin variants A and B in *Escherichia coli* for NMR studies. Protein Expression and Purification 70, 283–289.

Pérez, MD, Calvo, M, 1995. Interaction of β-lactoglobulin with retinol and fatty acids and its role as a possible biological function for this protein: a review. Journal of Dairy Science 78, 978–988.

Pérez, M.D., Puyol, P., Ena, J.M., Calvo, M., 1993. Comparison of the ability to bind lipids of beta-lactoglobulin and serum albumin of milk from ruminant and non-ruminant species. Journal of Dairy Research 60, 55–63.

Pérez, M.D., Sanchez, L., Aranda, P., Ena, J.M., Oria, R., Calvo, M., 1992. Effect of beta-lactoglobulin on the activity of pregastric lipase. A possible role for this protein in ruminant milk. Biochimica et Biophysica Acta 1123, 151–5.

Qin, B., Creamer, L., Baker, E., Jameson, G., 1998. 12-Bromododecanoic acid binds inside the calyx of bovine β-lactoglobulin. FEBS Letters 438, 272–278.

Reddy, I.M., Kella, N.K., Kinsella, J.E., 1988. Structural and Conformational Basis of the Resistance of B-Lactoglobulin to Peptic and Chymotryptic Digestion. Journal of Agricultural Food Chemistry 36, 737–741.

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Research 42, 320–324.

Roefs, S., Kruif, K., 1994. A Model for the Denaturation and Aggregation of β-Lactoglobulin. European Journal of Biochemistry 226, 883–889.

Saarinen, K.M., Pelkonen, A.S., Mäkelä, M.J., Savilahti, E., 2005. Clinical course and prognosis of cow's milk allergy are dependent on milk-specific IgE status. The Journal of allergy and clinical immunology 116, 869–75.

Sawyer, L., Kontopidis, G., 2000. The core lipocalin, bovine β-lactoglobulin. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1482, 136–148.

Seppälä, M., Taylor, R., Koistinen, H., Koistinen, R., Milgrom, E., 2002. Glycodelin: a major lipocalin protein of the reproductive axis with diverse actions in cell recognition and differentiation. Endocrine Reviews 23, 401–30.

Sievers, F., Higgins, D.G., 2013. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. Multiple Sequence Alignment Methods 1079, 105–116.

Uhrínová, S., Smith, M.H., Jameson, G.B., Uhrín, D., Sawyer, L., Barlow, P.N., 2000. Structural changes accompanying pH-induced dissociation of the beta-lactoglobulin dimer. Biochemistry 39, 3565–74.

Wang, C., Zhu, Y., Wang, J., 2015. Comparative study on the heat stability of goat milk and cow milk. Indian Journal of Animal Research 50, 610–613.

Wu, S.Y., Pérez, M., Puyol, P., Sawyer, L., 1999. β-Lactoglobulin Binds Palmitate within Its Central Cavity. Journal of Biological Chemistry 274, 170–174.

Yang, M., Guan, H., Liu, M., Lin, Y., Yang, J., Chen, W., Chen, C., Mao, S.J., 2008. Crystal structure of a secondary vitamin D3 binding site of milk β-lactoglobulin. Proteins: Structure, Function, and Bioinformatics 71, 1197–1210.

Zadow, J.G., Hardham, J.F., Kocak, H.R., Mayes, J.J., 1983. The stability of goat's milk to UHT processing. The Australian Journal of Dairy Technology 1, 20–23.

Chapter Two

2. Methods

2.1 Experimental reagents

2.1.1 Chemical reagents

Buffer and media components were purchased from Merck-Millipore, AppliChem, AnalaR, Thermofisher Scientific or Sigma-Aldrich. Urea was purchased from AnalaR. Equipment required for sodium-dodecyl-sulfate polyacrylamide-gel-electrophoresis (SDS-PAGE), including sample and running buffers, pre-cast gels and protein molecular weight ladders were purchased from Thermofisher Scientific. Antibiotics were purchased from AppliChem or Sigma-Aldrich. Sypro® Orange dye, SYBR® Safe DNA gel stain and fluorescein isothiocyanate (FITC) were purchased from Thermofisher Scientific. 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM) dye, dithiothreitol (DTT) and dimethylformamide (DMF) were purchased from Sigma-Aldrich. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Merck-Millipore.

2.1.2 Biological reagents

Prof. Geoff Jameson (Massey University) kindly provided the pETDuet plasmids containing bovine β-lactoglobulin (βlg) A and bovine βlg B genes, and *Escherichia coli* Origami (DE3) cells. Synthetic genes were purchased from Epoch Life Sciences. Chemically competent *E. coli* XL-1 Blue and BL21 (DE3) cells were purchased from Agilent Technologies and New England Biolabs. Restriction enzymes were purchased from New England Biolabs. Ligase was purchased from

Takara. Bovine and caprine immunoglobulin G (IgG) protein was purchased from Jackson Immunoresearch. Fresh cow milk (Holstein breed) was obtained from the Fairchild Dairy Teaching and Research Centre at the University of New Hampshire, United States of America, and fresh goat milk from a local farmer in either New Hampshire, United States of America (Nubian breed) or Canterbury, New Zealand (Nubian/Saanen cross breed).

2.2 Microbiology

2.2.1 Aseptic technique

Any media or glassware intended for use in bacterial culture was autoclaved at 121°C for 20 minutes prior to use. Loops and glass spreaders were sterilised by passing through a Bunsen flame before their use in inoculating bacteria. Antibiotics were filter sterilised by passing through a 0.2 µm syringe filter. All bacterial culture procedures were conducted near a blue Bunsen flame or in a laminar flow hood whilst wearing gloves.

2.2.2 Media

Luria Bertani (LB) broth was prepared from LB base powder by dissolving 20 g of powder per 1 L of distilled water before being autoclaved and stored at room temperature. If required, agar was added to LB to a final concentration of 15% (w/v) then autoclaved. When the temperature of the media returned to 50°C the necessary antibiotics were added. The agar was poured into sterile Petri plates and left to set before being stored at 4°C.

2.2.3 Bacterial strains

Three strains of *E. coli* were used in this study. *E. coli* XL1-Blue cells were used as a plasmid storage strain, while *E. coli* BL21 (DE3) and *E. coli* Origami (DE3) were used as protein expression strains (Table 2.1).

Table 2.1: Bacterial strains and genotypes.

E. coli strain	Genotype
XL1-Blue	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F´ *pro*AB *lac*IqZΔM15 Tn10 (Tet^r)]
BL21 (DE3)	F^- *ompT hsd*S_B(r_B^- m_B^-) *gal dcm* (DE3)
Origami (DE3)	F^- *omp*T *hsd*S_B(r_B^- m_B^-) *gal dcm lacY1 ahpC* (DE3) *gor522*:: Tn*10 trxB* (Kan^R, Tet^R)

2.2.4 Antibiotics

Antibiotics were prepared as 1000x stock solutions and stored at -20°C. The working concentrations used in this thesis are listed in Table 2.2.

Table 2.2: Stock and working concentrations of antibiotics used in this thesis.

Antibiotic	Stock Concentration	Working Concentration	Solvent
Kanamycin	15 mg/mL	15 µg/mL	Water
Ampicillin	100 mg/mL	100 µg/mL	Water
Tetracycline	12.5 mg/mL	12.5 µg/mL	Methanol

2.2.5 Preparation of competent cells

E. coli cells were streaked onto agar plates containing selective antibiotics and grown overnight at 37°C. A single colony from the agar plate was inoculated into 10 mL of LB and incubated overnight at 37°C with shaking. 250 µL of this pre-culture was inoculated into 2x 25 mL LB and incubated at 37°C with shaking. When the optical density at 600 nm (OD_{600}) reached 0.5 the

cells were placed on ice. The cells were centrifuged at 3000 g for 5 minutes at 4°C. The supernatant was discarded and tubes inverted on paper towels to remove any residual media. The pellet was gently resuspended in 30 mL of ice cold 0.1 M $CaCl_2$ using a pipette. Cells were incubated on ice for 30 minutes. Cells were centrifuged for a further 5 minutes at 3000 g and 4°C. The supernatant was removed and the pellet resuspended in 2 mL of ice cold calcium chloride (0.1 M). 500 µL of sterile 80% glycerol (v/v) was added. 100 µL aliquots of cells in sterile Eppendorf tubes were frozen in liquid nitrogen and stored at -80°C.

2.2.6 Transformation of competent cells

A 50 µL aliquot of competent cells was incubated on ice for 20 minutes with 2 µL of plasmid deoxyribose nucleic acid (DNA). The cells were heat shocked at 42°C for 90 seconds then returned to ice for 5 minutes. LB (400 µL) was added to the cells, which were then incubated for 1 hour with shaking at 37°C. The mixture (200 µL) was spread onto antibiotic-containing agar plates with a sterile spreader. Plates were incubated overnight at 37°C and cell growth in the presence of selective antibiotics indicated successful transformation. Untransformed competent cells were also streaked onto antibiotic containing plates to ensure that no contaminating plasmids were present.

2.2.7 Bacterial cultures

Pre-cultures of *E. coli* were prepared in sterile glass conical flasks or Falcon tubes by inoculating 5 – 50 mL of autoclaved LB, containing selective antibiotics, with a single colony from an agar plate and incubating overnight at 37°C with shaking. Large cultures (400 mL – 1 L), containing the same antibiotics, were inoculated with a 1/100 dilution of the pre-culture and grown at 37°C with shaking, in large baffled conical flasks. Culture growth was estimated by monitoring the OD_{600}.

2.2.8 Glycerol stocks

Bacterial strains were stored as glycerol stocks. For preparation of glycerol stocks a single colony was selected from an agar plate and inoculated into 50 mL of LB, containing selective antibiotics. The culture was grown overnight at 37°C with shaking. Aliquots (600 µL) were added to 200 µL sterile glycerol (80% v/v) in Eppendorf tubes, frozen in liquid nitrogen and stored at -80°C.

2.3 Molecular Biology

2.3.1 Plasmid isolation from transformed *E. coli*

Plasmids of interest were purified from XL1-Blue *E. coli* following an overnight pre-culture using an Invitrogen 'Quick plasmid mini-prep kit' according to the manufacturer's instructions. Typical yields of 75 µL of DNA at 30-80 ng/µL were obtained. The concentration of DNA was estimated using a NanoDrop® ND-1000 Spectrophotometer. DNA was stored at -20°C.

2.3.2 Restriction enzyme digestion

Restriction enzyme digestion reactions were prepared in Eppendorf tubes as described in Table 2.3. The reactions were incubated at 37°C for 3 hours. Digestion products were purified using agarose gel electrophoresis (see section 2.3.3).

2.3.3 Agarose gel electrophoresis

A 1% agarose gel was prepared by adding 0.5 g agarose to 50 mL TAE (Tris-acetate-ethylene-diamine-tetraacetic-acid) buffer along with 2.5 µL SYBR® Safe DNA gel stain. The solution was microwaved until the agarose dissolved, then poured into a gel base and left to set in the fridge. Once set, the comb was removed and TAE buffer poured on top of the gel. DNA gel loading dye (5 µL) was added to each DNA sample before loading into separate wells. A 1kb HyperLadder™ was loaded into the first well. The gel was run at 120 V for 40 minutes. Bands

Table 2.3: Restriction enzyme digestion reaction.

Reagent	Volume (µL)
CutSmart® Buffer	2
Plasmid DNA	5
Restriction enzyme 1	0.75
Restriction enzyme 2	0.75
Water	11.5
Total	20

were visualised under ultra-violet light and excised from the gel using a razor blade. Restriction digest products purified by agarose gel electrophoresis were extracted from agarose gel using a Roche 'Agarose Gel Extraction Kit', following the manufacturer's instructions.

2.3.4 Ligation

Ligation reactions were prepared in Eppendorf tubes as described in Table 2.4. The reactions were incubated at room temperature for 15 minutes. The product of the ligation reaction was used directly to transform competent *E. coli* cells.

2.3.5 Sequencing

Isolated plasmids were sent to Macrogen (Korea) for dideoxynucleotide sequencing to confirm that molecular cloning techniques had been successful. The sequencing results were analysed by aligning the reverse complement sequences against those expected using CLC Sequence Viewer (Qiagen Bioinformatics).

Table 2.4: Ligation reaction

Reagent	Volume (µL)
Vector	1
Insert	1.5
Ligase	2.5
Total	5

2.3.6 Bioinformatic analyses

The basic local alignment search tool (BLAST) (Altschul *et al*., 1990) was used to ascertain the level of nucleotide and amino acid similarity between DNA and protein sequences. Clustal Omega (Sievers and Higgins, 2013) was used to produce multiple sequence alignments.

2.4 Protein Expression

2.4.1 Small-scale protein expression

Small-scale protein expression trials were carried out to assess the best conditions for producing maximal amounts of soluble protein. *E. coli* BL21 (DE3) or Origami (DE3) cells containing the plasmid of interest were streaked onto agar plates containing selective antibiotics. A single colony from the agar plate was inoculated into 25 mL of LB containing the same antibiotics. The pre-culture was incubated at 37°C with shaking overnight. A small volume of the pre-culture was inoculated into 100 mL of LB containing appropriate antibiotics. The cultures were grown at 37°C with shaking until the OD_{600} was 0.4 - 0.7.

Protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM. The flasks either remained at 37°C or were moved to 26°C. Samples were taken 3 hours post-induction and after overnight expression. Cells within samples were harvested by centrifugation at 8000 g for 15 minutes then re-suspended in 20 mM Bis-tris, pH 6.5. Cells were

disrupted by sonication (Heilscher UP200S Ultrasonic Processor) in 0.5 seconds on, 0.5 seconds off bursts at 70% amplitude for 5 minutes on ice. Cell debris was pelleted by centrifugation at 20,000 g for 30 minutes. Samples of the supernatant and the pellet were analysed by SDS-PAGE (see section 2.5.4) to ascertain what proportion of the over-expressed protein was produced in a soluble form.

2.4.2 Large scale protein expression

A pre-culture was prepared as in section 2.4.1. The pre-culture was inoculated into 400 mL of LB containing the appropriate antibiotics. The culture was grown until OD_{600} = 0.4 - 0.6 at which point expression was induced by the addition of IPTG. The cultures were incubated overnight at 26°C with shaking. The cells were harvested by centrifugation at 8000 g for 15 minutes and the cell pellets were stored at -20°C until needed.

2.5 Protein Purification

The following purification protocol was utilised to purify the bovine βlg A and B variants and caprine βlg. The protocol was adapted from (Ponniah *et al.*, 2010) .

2.5.1 Anion exchange chromatography

The cell pellet (as obtained in section 2.4.2) was thawed on ice and resuspended in 20 mL buffer (20 mM Bis-tris, pH 6.5). The cells were disrupted by sonication for 6 x 5 minutes in 0.5 second pulses at 70% amplitude. The cell debris was pelleted by centrifugation at 20,000 g for 30 minutes and the lysate was loaded onto an anion exchange column (GE Healthcare) pre-equilibrated with buffer (20 mM Bis-tris, pH 6.5). Protein was eluted with a linear gradient of 0-1 M NaCl and 1.5 mL fractions were collected. Fractions were analysed by SDS-PAGE and those containing βlg were pooled.

2.5.2 Salt fractionation

The protein concentration of the pooled anion exchange fractions, estimated by absorbance at 280 nm, was diluted to 2 mg/mL. The pH of the solution was adjusted to 2.6 by drop wise addition of HCl with stirring. NaCl was added to the solution to a final concentration of 7% (w/v) and the sample was stirred at room temperature for 30 minutes. Precipitated *E. coli* proteins were removed by centrifugation at 30,000 g for 20 minutes. NaCl was added to the recovered supernatant to a final concentration of 30% (w/v) and stirred for 30 minutes to precipitate βlg. The solution was re-centrifuged at 30,000 g for 20 minutes and the supernatant was discarded. The pellet was resuspended in a small volume of buffer (50 mM Tris, 50 mM NaCl, pH 7.5).

2.5.3 Size exclusion chromatography

The solubilised βlg protein was injected onto a size exclusion column (HiLoad Superdex 200 16/60 120 mL) equilibrated with buffer (50 mM Tris, 50 mM NaCl, pH 7.5). The protein was eluted using a flow rate of 1 mL/min. Fractions were analysed by SDS-PAGE and those containing pure βlg were pooled. The protein was concentrated to at least 10 mg/mL using a molecular-weight cut-off spin filter (Sartorius). Aliquots (200 µL) of protein were frozen in liquid nitrogen and stored at -80°C.

2.5.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Samples for SDS-PAGE were prepared in Eppendorf tubes as described in Table 2.5. A precast BOLT™ 4-12% Bis-Tris gel was placed into a gel chamber and immersed in Bolt™ 2-(N-morpholino)ethanesulfonic acid (MES) SDS running buffer. Novex® Sharp pre-stained protein standard (5 µL) was loaded into the first well of the gel, followed by 10 µL of each sample into the remaining wells. Electrophoresis was conducted at 165 V for 35 minutes. The gel was then removed from its casing and stained with SimplyBlue™ SafeStain.

Table 2.5: SDS-PAGE sample preparation.

Reagent	Volume (µL)
Protein sample	5
Novex® lithium dodecyl sulfate sample buffer (4x)	2.5
DTT (500 mM)	1
Water	1.5
Total	10

2.5.5 Protein quantification

The absorbance of protein solutions at 280 nm was measured using a Cary UV/Vis spectrophotometer or a Nano-Drop ND-1000 spectrophotometer. This was used to calculate protein concentration using the extinction coefficients at 280 nm, calculated from amino acid sequences using ExPASy ProtParam (Gasteiger *et al.*, 2005) (Table 2.6), and Beer's law.

Table 2.6: β-Lg extinction coefficient parameters.

Protein	Extinction Coefficient ($M^{-1}cm^{-1}$)
Bovine βlg A	17210
Bovine βlg B	17210
Caprine βlg	17210

2.5.6 Theoretical isoelectric point estimations

The theoretical isoelectric points (i.e. the pH at which a protein has no net charge) of bovine and caprine βlg proteins were estimated from amino acid sequences using ExPASy ProtParam (Gasteiger *et al*., 2005) and SEDNTERP (Laue *et al*., 1992) .

2.6 Biophysical and Physicochemical Characterisation Methods

2.6.1 Mass spectrometry

Protein samples to be analysed by mass spectrometry were exchanged into water using a 5 mL desalting column (GE Healthcare) and adjusted to 1 mg/mL. They were analysed using a maXis 3G Ultrahigh Resolution Quadrupole Time-of-Flight mass spectrometer, equipped with an electrospray ionisation source (Bruker Daltonik GmbH, Bremen, Germany).

2.6.2 Denaturation experiments

2.6.2.1 Circular dichroism spectroscopy

Circular dichroism (CD) spectra were collected using a Jasco J-815 CD spectrometer. Far-UV spectra were collected for protein samples at a concentration of 0.1 mg/mL using a 2 mm quartz cuvette between 180 nm and 260 nm. Near-UV spectra were collected for protein samples at a concentration of 0.5 mg/mL using a 1 cm quartz cuvette between 250 nm and 320 nm. The recorded protein spectra were blanked against buffer spectra and any data collected which exceeded a high tension (HT) voltage of 600 were removed.

The resulting ellipticity data were converted to mean residue ellipticity (MRE) by multiplying the data by a conversion factor (equations 2.1 and 2.2).

Equation 2.1:

$$Mean\ residue\ weight\ (MRW) = \frac{Molecular\ weight}{Number\ of\ amino\ acids - 1}$$

Equation 2.2

$$Conversion\ factor = \frac{MRW}{Path\ length(mm)\ x\ [sample](mg/ml)}$$

2.6.2.1.1 Heat denaturation

Thermal denaturation of protein samples was monitored by recording the ellipticity at 293 nm with data collected at 1°C intervals between 25 and 95°C. Proteins were analysed at a concentration of 0.5 mg/mL in a 1 cm quartz cuvette and were stirred with a small magnetic stirrer bar. A Boltzmann sigmoid curve was fitted to the data in Origin (OriginLab, Northampton, MA).

2.6.2.1.2 Urea denaturation

Urea denaturation of protein samples was monitored at 218 nm at concentrations of urea between 0 and 9 M. Two stock solutions were prepared, one containing 0.2 mg/mL protein, and the other containing 0.2 mg/mL of protein and 9 M urea. Different quantities of these stock solutions were mixed to create a dilution range of urea while maintaining a constant protein concentration. Samples were left at room temperature for 4 hours before the ellipticity at 218 nm was measured. Samples were analysed in 2 mm quartz cuvettes. A Boltzmann sigmoid curve was fitted to the data in Origin (OriginLab, Northampton, MA).

Denaturation reversibility was assessed by diluting a urea-denatured protein sample and then re-measuring the spectra. Three samples of protein (caprine βlg) were prepared, one at 0.2 mg/mL protein and 1 M urea, the second at 0.2 mg/mL protein and 9 M urea and the third at 1.8 mg/mL protein and 9 M urea. The samples were left for four hours, then CD spectra were recorded for samples 1 and 2. Sample 3 was diluted with buffer so that the final protein concentration was 0.2 mg/mL and the final urea concentration was 1 M urea. All three samples were left overnight and then their CD spectra recorded.

2.6.2.2 Differential scanning fluorimetry

Two methods were used to conduct differential scanning fluorimetry.

The first method utilised the BioRad iQ5 Real-Time polymerase-chain-reaction (PCR) Detection System and the dye Sypro® Orange. Protein (80 µL) at a concentration of 0.5 mg/mL was mixed with 20 µL of 50x concentrated Sypro® Orange dye and added to the wells of a 96-well PCR plate. The plates were sealed then placed in an iQ5 Real-Time PCR Detection System. The samples were heated from 20 to 95°C with fluorescence measurements made every 0.2°C.

The second method utilised a Cary Eclipse spectrofluorometer and the dye 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM). CPM dye was dissolved in DMF at 4 mg/mL and stored at -20°C. Immediately prior to use the dye was diluted 1/40 with buffer. 3 mL samples were prepared in quartz cuvettes containing 100 µL of dye and a final protein concentration of 0.05 mg/mL (~2.5 µM). The cuvettes (up to four at a time) were placed in the spectrofluorometer and heated from 20 – 100°C at 1°C per minute, with fluorescence measurements made every 0.5°C. The excitation and emission wavelengths were set at 387 and 463 nm, respectively. A Boltzmann sigmoid curve was fitted to the data in Origin (OriginLab, Northampton, MA).

2.6.3 Analytical ultracentrifugation

2.6.3.1 Sedimentation velocity analysis

Sedimentation velocity experiments were conducted in a Beckman Coulter XL-I analytical ultracentrifuge. Sedimentation was monitored utilising one of the three available optical systems (absorbance, interference and fluorescence). All three optical systems were used during the course of this work. Experiments conducted using fluorescence-detection optical system were carried out at the University of New Hampshire, United States of America.

Reference solution (400 µL) and sample solutions (380 µL) were added to 12 mm double sector cells with quartz or sapphire windows. Cells were mounted in an An-50 Titanium eight-hole rotor. Initial scans were performed at 3,000 rpm to determine the optimal settings, then data were collected in continuous mode at 50,000 rpm and 20 or 25°C.

Buffer density and viscosity and an estimate of the partial specific volume of proteins were calculated using SEDNTERP (Laue *et al.*, 1992) . Data were fitted to a continuous sedimentation coefficient [*c*(*s*)] model and a continuous mass [*c*(*M*)] model using SEDFIT (Schuck, 2000) .

2.6.3.2 Fluorescent labelling of proteins for interaction studies using fluorescence detection analytical ultracentrifugation

Proteins were labelled with fluorescein isothiocyanate (FITC) for detection using fluorescence optics by incubating proteins at 2 mg/mL (in 0.1 M sodium carbonate, pH 9) with 50 µL FITC per 1 mL of protein. FITC was dissolved in DMF at 1mg/mL immediately prior to use. The solution was protected from light using tinfoil and rotated to mix for one hour at room temperature. The solution was passed through a 5 mL desalting column (GE Healthcare) to remove the bulk of the free label and further cleaned by gel filtration chromatography (HiLoad Superdex 200 16/60 120 mL). The degree of labelling was calculated by measuring the absorbance at 280 nm and 494 nm. The protein concentration was calculated from the absorbance at 280 nm taking into account the contribution of FITC as per equation 2.3.

Equation 2.3:

$$Protein\ concentration\ (M) = \frac{A_{280} - (A_{max}\ x\ CF)}{\varepsilon}$$

Where A_{max} is the wavelength of maximum absorbance for the dye molecule (for FITC this is 494 nm), CF is the correction factor which adjusts for the amount of absorbance at 280 nm caused by the dye (for FITC this 0.3, as supplied by Thermofisher Scientific), and ε is the protein molar extinction coefficient (for βlg this is 17210 M^{-1} cm^{-1}).

The degree of labelling was calculated as according to equation 2.4.

Equation 2.4:

$$Moles\ dye\ per\ mole\ protein = \frac{A_{max}\ of\ the\ labelled\ protein}{\varepsilon' x\ protein\ concentration\ (M)}$$

Where ε' is the molar extinction coefficient of the fluorescent dye (for FITC this 68000 M^{-1} cm^{-1}, as supplied by Thermofisher Scientific). A degree of labelling between 0.3 and 1.0 moles of dye per mole of protein was deemed acceptable.

2.6.3.3 Preparation of milk for interaction studies using fluorescence detection analytical ultracentrifugation

Fresh, raw, cow and goat milk was obtained (see section 2.1.2 for breeds and locations). Before use, the milk was spun at 5000 g for 5 minutes. The layer of fat that resulted at the top was removed and the liquid underneath was transferred to a new vessel, leaving behind the small pellet of cells that formed at the bottom of the milk sample. Diluted 'skimmed' milk samples were used for all interaction studies.

2.6.4 X-ray crystallography

2.6.4.1 Crystallisation

Sitting-drop vapour-diffusion crystallisation screens were performed using a Mosquito crystallisation robot (TTP Labtech) and the commercial crystallisation screens PACT, JCSG+, Clear Strategy Screens 1 and 2, Midas and Morpheus (Molecular Dimensions Ltd). Drops consisted of 400 nL protein solution (10 mg/mL in 50 mM Tris, 50 mM NaCl, pH 7.5) and 400 nL reservoir solution, with wells containing 100 µL reservoir solution. Crystal trays were incubated at room temperature (~20°C).

2.6.4.2 Data collection

The crystal was harvested with a litho-loop and flash-cooled in liquid nitrogen. Diffraction data were collected at 110 K (-163°C) on the MX2 beamline at the Australian Synchrotron, Victoria, Australia, at a wavelength of 0.9537 Å. The detector was positioned 125 mm from the crystal and data were collected in 1° steps over a 360° pass with an exposure time of 0.5 seconds each.

2.6.4.3 Structure determination

Diffraction images were indexed and integrated using *iMOSFLM* (Battye *et al*., 2011) . Scaling and data reduction were performed using *SCALA* from the *CCP4* program suite (Winn *et al*., 2011) . Molecular replacement was performed by *MolRep* (Vagin and Teplyakov, 2010) using the crystal structure of bovine βlg A (PDB: 1BSO) as the search model. The structure was refined anisotropically using *REFMAC* (Murshudov *et al*., 1997) and *PHENIX* (Adams *et al*., 2010) and iterative rounds of model building were performed in *Coot* (Emsley *et al*., 2010) . The structure of cβlg was deposited in the protein data bank (PDB: 4TLJ).

2.6.5 Small-angle X-ray scattering

2.6.5.1 Data collection

Small-angle X-ray scattering (SAXS) data were collected at the Australian Synchrotron SAXS/WAXS beamline equipped with a Pilatus 1M detector. The wavelength of the X-rays was 1.0332 Å. The scattering vector, q, is defined as $4\pi\sin(\theta)/\lambda$. The sample-to-detector distance was 1600 mm, which provided a q range of 0.01–0.55 $Å^{-1}$. The protein was subjected to size-exclusion chromatography (Superdex 200 5/150 GL column) prior to analysis to remove any aggregates, and flowed through a 1.5 mm diameter thin-walled glass capillary during data collection to limit exposure of the sample to radiation, limiting the risk of aggregation and degradation. Proteins were analysed at a starting concentration of 10 mg/mL in 50 mM sodium phosphate, 50 mM sodium chloride, pH 7.5.

2.6.5.2 Data analysis

Two-dimensional scattering intensity images were radially averaged and background-subtracted using *SCATTERBRAIN* (Australian Synchrotron). Data analyses were performed using the *ATSAS* 2.7 package (Petoukhov *et al*., 2012). Guinier fits were produced using *PRIMUS* (Konarev *et al*., 2003). Indirect Fourier transforms were performed using *GNOM* (Svergun, 1992) to yield the function $P(r)$, which gives the relative probabilities of distances, r, between scattering centres, and the maximum dimension of the particle, D_{max}. Theoretical SAXS

scattering curves of crystal structures were generated from atomic coordinates and compared with the experimentally determined scattering curves using *CRYSOL* (Svergun *et al*., 1995).

2.6.6 Molecular dynamics simulations

2.6.6.1 Simulation set-up and procedure

Molecular dynamics simulations were carried out on dimers of bovine βlg A (PDB: 1BSO or 1BSY), bovine βlg B (PDB: 1BSQ) and caprine βlg (PDB: 4TLJ) using GROMACS version 4.0 or version 4.6.1. The force field selected for modelling the system was GROMOS 54A7 (Schmid *et al*., 2011) using the simple-point-charge (SPC) model (Berendsen *et al*., 1981) for water molecules. Missing side chain residues were modelled using PDB_Hydro (Azuara *et al*., 2006). The completed PDB format files were converted into a GROMACS compatible format using the pdb2gmx program within the GROMACS suite of programmes (Berendsen *et al*., 1995; Lindahl *et al*., 2001).

Completed structures were solvated in a cubic box with a minimum distance of 8 Å from the protein to the side of the box. NaCl (150 mM) was added to the aqueous phase. The pH of the system was neutral (pH 7). The resulting system was subjected to energy minimisation in order to remove any steric clashes and inappropriate geometry. The system was heated to 310.15 K (37°C), while applying position restraints to the protein atoms with a force constant k = 1000 $kJ.mol^{-1}$. Bond lengths and solvent bond angle positions were constrained using the LINCS algorithm (Hess *et al*., 1997), allowing an integration step of 2 fs with snapshots saved every 20 ps. Simulations were carried out under periodic boundary conditions. The temperature was maintained at 310.15 K using the Berendsen thermostat (Berendsen *et al*., 1984) with a temperature coupling time τ_T = 0.1 ps^{-1}. Subsequently the restraints were removed and simulations of up to 200 ns were performed under normal pressure and temperature (NPT) conditions. The pressure was maintained at 1 atm using the Berendsen barostat with a pressure coupling constant of τ_P = 0.5 ps^{-1}.

Simulations for each structure were carried out in segments of 10 – 50 ns at a time. They were then combined into a single trajectory using the GROMACS program 'trjcat'. In order to visualise the course of the simulations in Visual Molecular Dynamics (VMD) (version 1.9.2)

(Humphrey *et al*. 1996) the trajectories were processed using the GROMACS program 'trjconv' under the following conditions: -pbc mol -ur compact -center, which centres the solute molecules within the simulation box, while keeping the solute molecules whole and accounting for the periodic boundary conditions.

2.6.6.2 Simulation analysis

Analyses were performed using programs within the GROMACS suite:

G_rms was used to perform a least-squares fit between the Cα atoms of each structure along the trajectory and the starting structure, then to calculate the root mean square deviation from the initial structure of all atoms within the protein. To separate out internal structural changes from relative motions of one chain of the protein with respect to the other this can be carried out on the whole protein or for a single chain at a time.

G_rmsf was used to perform a least squares fit between the Cα atoms of each structure and the starting structure, then to calculate the root mean square fluctuation of each residue with respect to its average positon during the simulation.

G_saltbr computes the distances between all combinations of charged groups of atoms as a function of time. A distance cut-off of 0.4 nm was used such that groups never closer than that distance are not considered. A python script was used (supplied by Jane Allison) to select only those interactions that occur between chain A and chain B.

G_hbond computes hydrogen bonds that occur between two groups of atoms, in this case chain A and chain B of each protein. A python script was used (supplied by Jane Allison) that produces a time-series of the hydrogen bonds that occur for at least 20 ns during the simulation.

2.7 References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D 66, 213–221.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Azuara, C., Lindahl, E., Koehl, P., Orland, H., Delarue, M., 2006. PDB_Hydro: incorporating dipolar solvents with variable density in the Poisson-Boltzmann treatment of macromolecule electrostatics. Nucleic acids research 34, 38–42.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G., 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallographica Section D 67, 271–281.

Berendsen, H., Postma, J., Gunsteren, W., DiNola, A., Haak, J., 1984. Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics 81, 3684–3690.

Berendsen, H.J., Postma, J.P., van Gunsteren, W.F., Hermans, J., 1981. Interaction Models for Water in Relation to Protein Hydration. D. Reidel Publishing Company, pp. 331–342.

Berendsen, H.J., van der Spoel, D., van Drunen, R., 1995. GROMACS: A message-passing parallel molecular dynamics implementation. Computer Physics Communications 91, 43–56.

Emsley, P., W, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallographica Section D 66, 486–501.

Gasteiger, E., Hoogland, C., Gattiker, A., S, Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server. Humana Press Inc., Totawa, NJ.

Hess, B., Bekker, H., Berendsen, H., 1997. LINCS: a linear constraint solver for molecular simulations. Journal of computational chemistry 18, 1463–1472.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H., Svergun, D.I., 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography 36, 1277–1282.

Laue, T.M., Shah, B.D., Ridgeway, T.M., Pelletier, S.L., 1992. Computer-aided interpretation of analytical sedimentation data for proteins. The Royal Society of Chemistry, Cambridge, UK, pp. 90–125.

Lindahl, E., Hess, B., van der Spoel, D., 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. Molecular Modeling 7, 306–317.

Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallographica Section D 53, 240–255.

Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Svergun, D.I., 2012. New developments in the ATSAS program package for small-angle scattering data analysis. Journal of Applied Crystallography 45, 342–350.

Ponniah, K., Loo, T., Edwards, P., Pascal, S., Jameson, G., Norris, G., 2010. The production of soluble and correctly folded recombinant bovine β-lactoglobulin variants A and B in *Escherichia coli* for NMR studies. Protein Expression and Purification 70, 283–289.

Schmid, N., Eichenberger, A.P., Choutko, A., Riniker, S., Winger, M., Mark, A.E., van Gunsteren, W.F., 2011. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. European Biophysics Journal 40, 843–856.

Schuck, P., 2000. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. Biophysical Journal 78, 1606–1619.

Sievers, F., Higgins, D.G., 2013. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. Multiple Sequence Alignment Methods 1079, 105–116.

Svergun, D.I., 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. Journal of Applied Crystallography 25, 495–503.

Svergun, D.I., Barberato, C., Koch, M.H., 1995. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates Journal of Applied Crystallography 28, 768–773.

Vagin, A., Teplyakov, A., 2010. Molecular replacement with MOLREP. Acta Crystallographica Section D 66, 22–25.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., P, Wilson, K.S., 2011. Overview of the CCP4 suite and current developmetns. Acta Crystallographica Section D 67, 235–242.

Chapter Three

3. Cloning, Expression and Purification of Bovine and Caprine β-Lactoglobulin

3.1 Introduction

3.1.1 Recombinant expression of bovine βlg

There have been many attempts made to produce native-like recombinant bovine βlg (Ariyaratne *et al*., 2002). Soluble bovine βlg has been expressed in mammalian cells, but not in large quantities (Hyttinnen *et al*., 1998), while production in yeast results in higher yields but the protein often has carbohydrate non-specifically associated with it (Denton *et al*., 1998; Kim *et al*., 1997). Bovine βlg expressed using the Baculovirus system in insect cells resulted in protein that was truncated at the C-terminus (Mizumachi *et al*., 1993). Early attempts to produce recombinant bovine βlg in *Escherichia coli* were unsuccessful at producing soluble protein (Batt *et al*., 1990; Chatel *et al*., 1996; Cho *et al*., 1994).

In 2002, Ariyaratne *et al*. (2002) successfully developed an *E. coli* expression system capable of producing soluble bovine βlg. Firstly, they designed an *E. coli* codon-optimised gene for bovine βlg A based on a codon-usage table for highly expressed *E. coli* proteins. They cloned the synthetic gene into a pTrxFus expression vector in order to express the protein as a thioredoxin fusion protein. They were then able to remove the thioredoxin via cleavage with enterokinase, leaving the native sequence plus three additional N-terminal residues. Unfortunately, this method did not yield consistent results and did not work well for the B variant of bovine βlg (Ponniah *et al*., 2010).

3.1.2 Disulfide bond formation is required for successful recombinant expression of βlg

The difficulty in consistently producing soluble bovine βlg lies in the fact that it contains two disulfide bonds that are critical for the correct folding (Ponniah *et al*., 2010). The cytoplasm within *E. coli* cells is not conducive to the formation of disulfide bonds due to its reducing nature. A mutant strain of *E. coli* has been developed (Bessette *et al*., 1999) in which the cytoplasmic redox potential is comparable to that of the mammalian endoplasmic reticulum, achieved through mutations in the thioredoxin reductase and glutathione reductase genes. These cells, known as Origami cells, allow disulfide bonds to form.

However, on their own these Origami cells may not be enough to produce correctly folded protein, particularly if the protein being expressed contains more than a single pair of cysteine residues, as non-native disulfide bonds may form. Co-expression with a chaperone protein, such as DsbC, a disulfide bond isomerase and chaperone, is often required to resolve these non-native disulfide bonds (Kurokawa *et al*., 2000). βLg contains five cysteine residues, four of which participate in disulfide bond formation. The simultaneous expression of bovine βlg and DsbC within Origami cells has proved to be a very effective method of producing soluble bovine βlg (Ponniah *et al*., 2010). Meanwhile, very few studies have focused on the caprine orthologue of βlg, and there have been no reports of the use of recombinant bacterial expression to produce this variant of the protein.

3.1.3 Overview

The overarching aim of this thesis is to characterise the physicochemical and biophysical properties of caprine βlg and to compare these to those of the bovine βlg variants. Large amounts of protein were desired in order to conduct a thorough characterisation of these variants of βlg. As non-heat-treated βlg does not contain any post-translational modifications (Fogliano *et al*., 1998), the choice was made to produce these proteins recombinantly as this would ensure that each protein was produced under very similar conditions and thus would be directly comparable.

Until now, it was unknown whether the caprine orthologue of βlg could be successfully recombinantly expressed and purified utilising the same protocols as those that have been

successful for the bovine variants of βlg. This chapter details the successful recombinant expression and purification of caprine βlg (and the bovine variants of βlg) by co-expressing these proteins with DsbC within *E. coli* Origami (DE3) cells.

Caprine βlg exhibits marked stability at low pH values, as does bovine βlg, which allows this protein to be purified to homogeneity by following the purification protocol for bovine βlg (Ponniah *et al*., 2010). This protocol involves the salting-out of proteins at low pH (less than pH 3), leaving behind pure βlg. Use of this protocol resulted in the successful purification of bovine βlg A and B and caprine βlg.

Mass spectrometry confirms that these recombinantly expressed βlg proteins (bovine A and B and caprine βlg) are successfully over-expressed as full-length constructs and that they contain the correct number of disulfide bonds. Circular dichroism spectra indicate that these proteins adopt a similar fold to each other and that this fold is indistinguishable from that of native βlg isolated from milk.

3.2 Results and discussion

3.2.1 Recombinant expression of bovine and caprine βlg

A synthetic gene construct was prepared in order to recombinantly express caprine βlg within *E. coli* Origami (DE3) cells. The caprine βlg gene was synthesised (Epoch Life Sciences) based on the codon-optimised sequence of the bovine βlg B gene. Six codons were altered in order to produce the six amino acid substitutions that are present between bovine βlg B and caprine βlg. Codons were chosen according to the codon preference of *E. coli*.

A pETDuet plasmid was utilised in order to co-express caprine βlg with DsbC (a disulfide bond isomerase). These plasmids harbour two multiple cloning sites, both under the control of a T7 promoter, which allow two proteins to be co-expressed upon induction with isopropyl β-D-1-thiogalactopyranoside (IPTG). Two pETDuet plasmids, containing the gene for DsbC in the first multiple cloning site and containing the gene for either bovine βlg A or bovine βlg B in the second, were supplied (Prof. Geoff Jameson). The bovine βlg genes are flanked by the Nde1 and Kpn1 restriction enzyme recognition sites. In order to produce a pETDuet-DsbC-cβlg

construct, the bovine βlg A gene was first removed from the pETDuet plasmid by digesting the plasmid using the restriction enzymes Nde1 and Kpn1 (section 2.3.2). The empty vector and the insert were separated by electrophoresis on an agarose gel (section 2.3.3). The synthetic gene for caprine βlg was removed from its pBlueScript vector using the same restriction enzymes and the open reading frame of caprine βlg was then ligated into the second multiple cloning site of the empty pETDuet vector (section 2.3.4).

3.2.1.1 Verification of correct construct

The pETDuet-DsbC-cβlg construct was used to transform competent *E. coli* XL1-Blue cells (section 2.2.6) and the cells were plated onto ampicillin-containing agar plates. The pETDuet vector confers ampicillin resistance; therefore, growth on ampicillin-containing plates indicated successful ligation and transformation. The empty pETDuet vector following the restriction digest was unable to confer ampicillin resistance to *E. coli* XL1-Blue cells indicating that complete digestion of the plasmid had occurred.

To confirm the identity of the gene insert, DNA was isolated from the *E. coli* XL1-Blue cells (section 2.3.1) and sequenced. The T7 reverse primer was used to sequence the second multiple cloning site of the pETDuet vector. In each case, sequencing confirmed that the correct insert was present in each plasmid. Figure 3.1 shows the protein sequences that arise from these genes.

Once the identity of the insert was confirmed, the construct was used to transform competent *E. coli* BL21 (DE3) and *E. coli* Origami (DE3) cells. Expression trials were carried out in both type of *E. coli* find the optimal conditions for producing soluble caprine βlg.

3.2.1.2 Expression trials using E. coli BL21 (DE3) cells

Recombinant expression of caprine βlg was initially attempted in *E. coli* BL21 (DE3) cells containing the pETDuet-DsbC-cβlg construct (section 2.4.1). Various conditions were assessed in order to identify the best conditions for producing maximal amounts of soluble protein. Small-scale cultures were grown to an optical density at 600 nm (OD600) of either

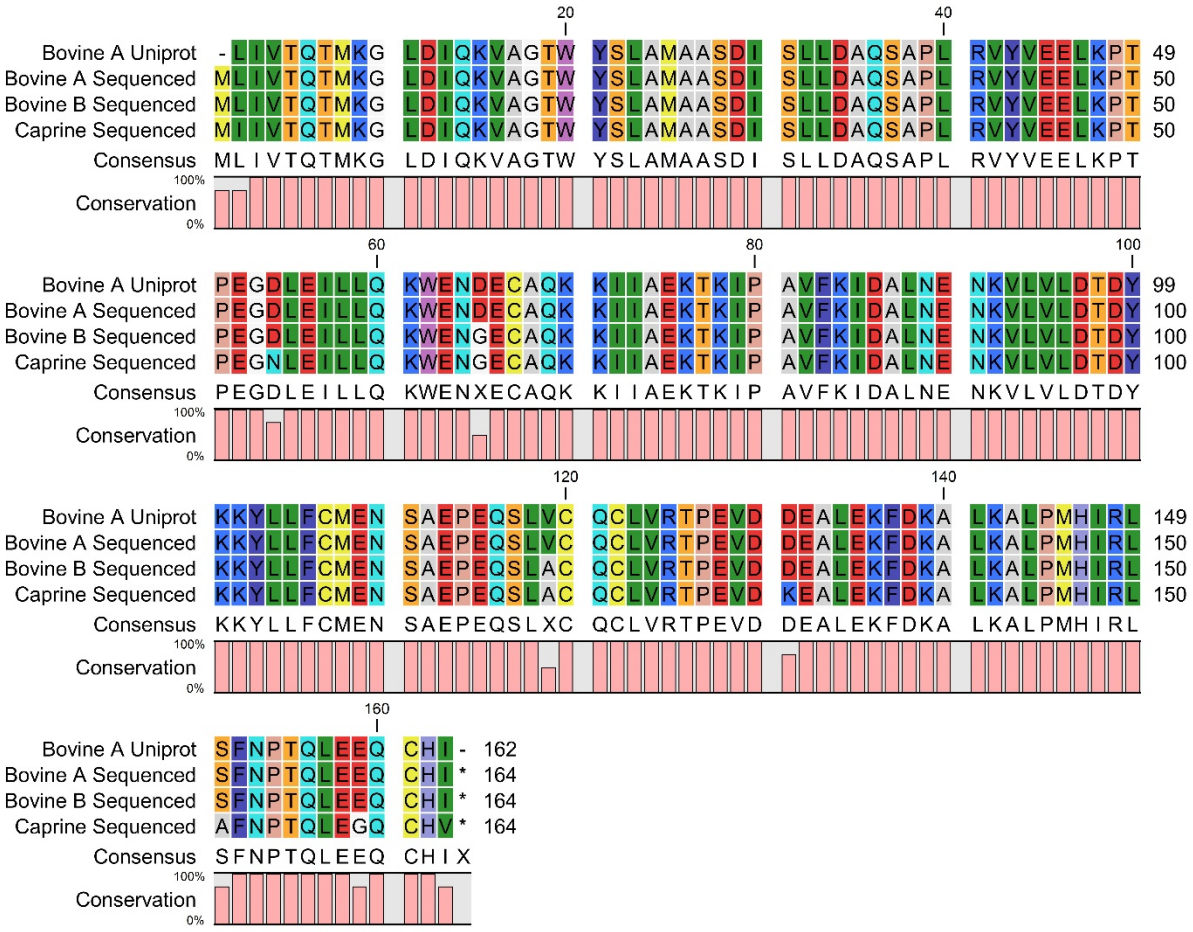


Figure 3.1: Protein sequences that arise from the bovine βlg A, bovine βlg B and caprine βlg genes (translated from DNA sequences using Expasy Translate). These sequences are aligned with the protein sequence of bovine βlg A obtained from the Uniprot database (Uniprot accession number PO2754). Alignments produced using CLC sequence viewer (Qiagen Bioinformatics).

0.5 or 0.7 before protein expression was induced with the addition of IPTG. The flasks either remained at 37°C or were moved to 26°C. Samples were taken 3 hours post-induction and after overnight expression.

The samples were sonicated and the cell debris pelleted by centrifugation. Samples of the supernatant and of the pellet were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (section 2.5.4) to assess what proportion of the over-expressed protein was produced in a soluble form. While a higher level of protein expression is seen after overnight expression than after 3 hours of expression, in all conditions the caprine βlg is present in the pellet (induction at OD600 = 0.5 or 0.7, expression for 3 hours or overnight, at

both 26°C and 37°C) (Figure 3.2). This indicates that expression in BL21 (DE3) cells is not suitable for producing soluble caprine βlg.

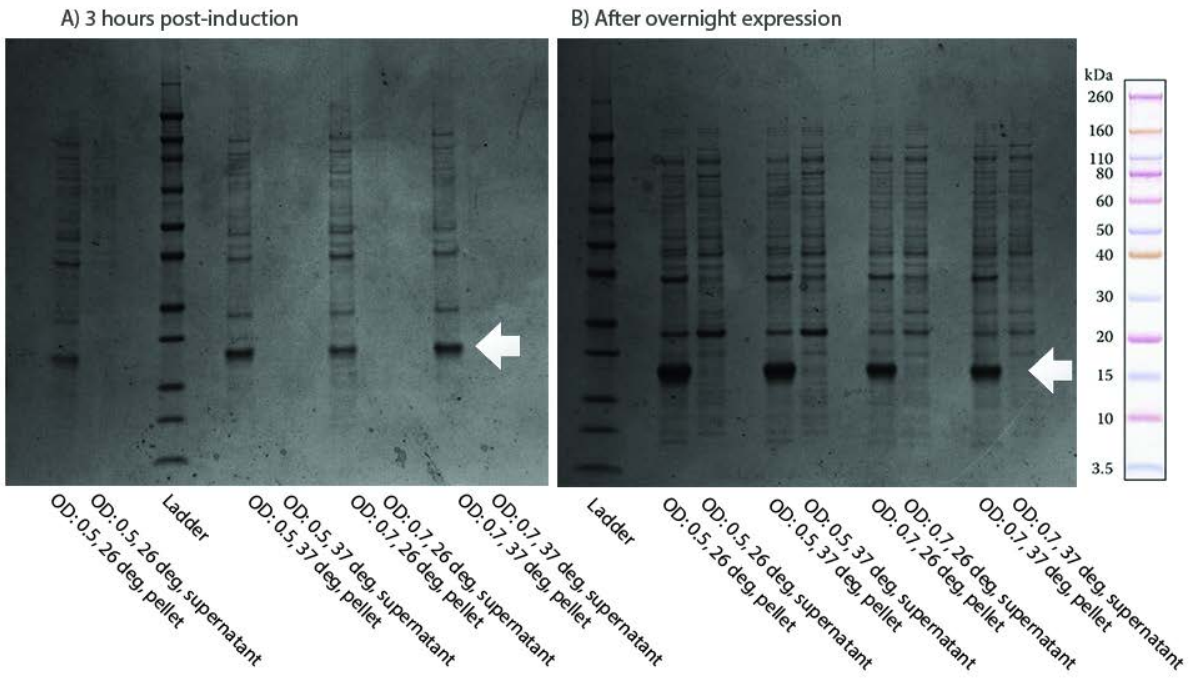


Figure 3.2: Expression trials of caprine βlg using *E. coli* BL21 (DE3) cells. The location of caprine βlg is indicated by an arrow (approximate molecular mass is 18 kDa).

3.2.1.3 Expression trials using E. coli Origami (DE3) cells

A similar expression trail was carried out using *E. coli* Origami (DE3) cells, except that expression was induced when the cultures reached an OD600 of 0.4. Expression was carried out at 26°C and 37°C and samples were taken after 3 hours and after overnight incubation. SDS-PAGE analysis (Figure 3.3) indicates that *E. coli* Origami (DE3) cells are far better suited to producing soluble caprine βlg than *E. coli* BL21 (DE3) cells. Caprine βlg is visible in the supernatant (and therefore soluble) when expressed overnight at either 26°C or 37°C. It is also present when expressed at 37°C for 3 hours but not when expressed at 26°C for 3 hours. This is most likely due to the lower rate of protein expression at this temperature. βLg is also visible in the pellet in each condition, indicating that only a fraction of the protein is soluble. However, expression overnight at 26°C resulted in the greatest proportion of caprine βlg in a soluble

form. Thus, this condition was chosen for future expression conditions for all three of the βlg proteins examined in this study (bovine βlg A and B and caprine βlg).

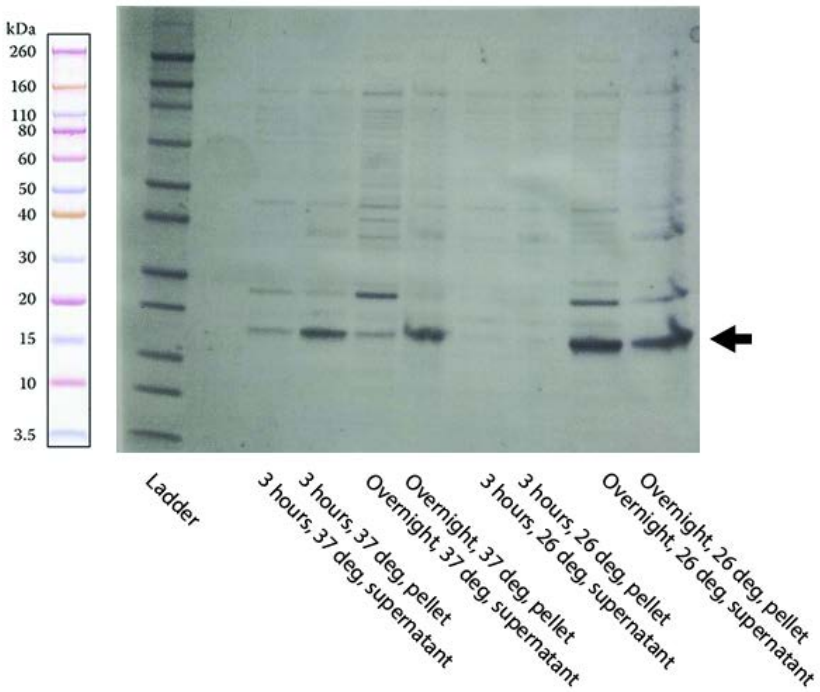


Figure 3.3: Expression trials of caprine βlg using *E. coli* Origami (DE3) cells. The location of caprine βlg is indicated by an arrow (approximate molecular mass is 18 kDa).

3.2.2 Purification of bovine and caprine βlg

Purification of each of the three βlg proteins was achieved by exploiting the acid stability exhibited by these proteins when correctly folded. This serves to remove not only any contaminating proteins but also any misfolded βlg proteins.

Firstly, the cells were lysed by sonication and the clarified lysate then subjected to anion-exchange chromatography (section 2.5.1). The protein elutes from the anion-exchange column at NaCl concentrations between 30 and 60%. This purification step removes many of the contaminating *E. coli* proteins (Figure 3.4). The considerable acid stability of both bovine and caprine βlg, when correctly folded, provides an effective second purification step (section 2.5.2) since at a pH of less than 3, and at moderate salt concentrations (7% NaCl), almost all of

the remaining contaminating proteins are precipitated. The remaining protein, which is predominantly βlg, can then be precipitated with high concentrations of salt (30% NaCl). After dissolving the precipitate in buffer, a final polishing by size-exclusion chromatography (section 2.5.3) results in very pure protein, as estimated by SDS-PAGE (Figure 3.4). This protocol yields around 20 mg of protein per litre of culture.

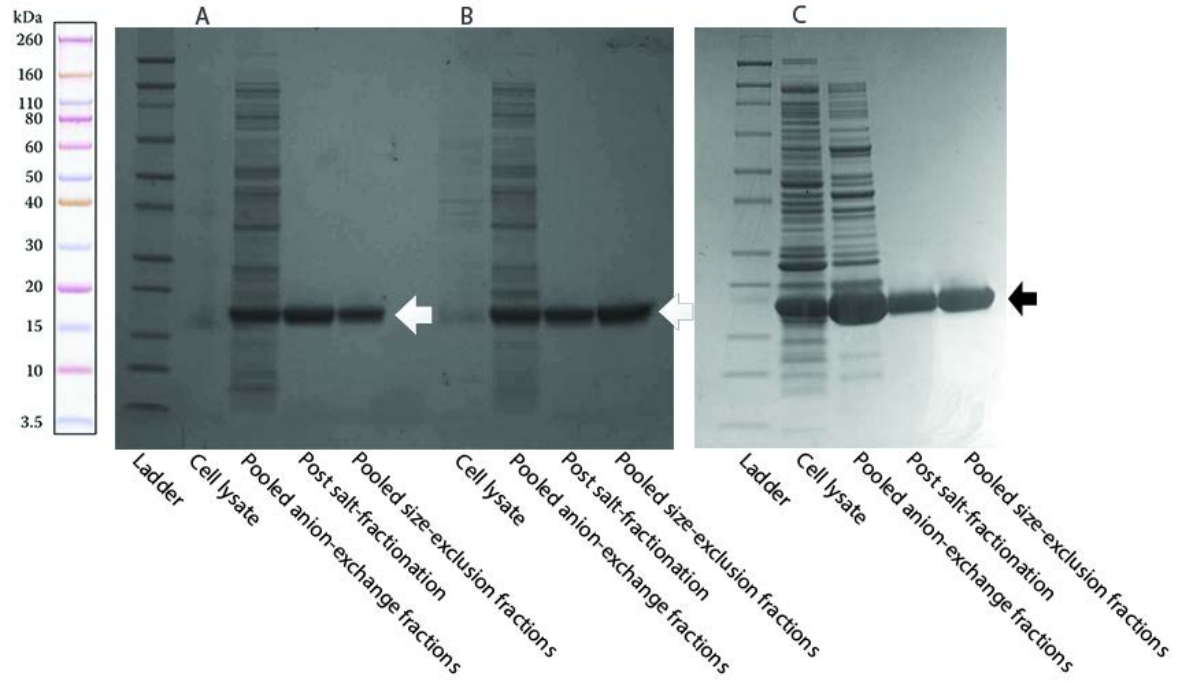


Figure 3.4: Purification gels of A) bovine βlg A, B) bovine βlg B, and C) caprine βlg. The location of βlg is indicated by the arrow.

3.2.3 Mass spectrometry confirms the correct protein products

Mass spectrometry confirms that all three proteins were expressed as full-length protein constructs. The molecular masses of the three βlg proteins examined in this study were estimated according to their amino acid sequences, using the ExPasy ProtParam tool (Gasteiger *et al.*, 2005), and are listed in Table 3.1. The deconvoluted mass values obtained by mass spectrometry, also listed in Table 3.1, are all 4 daltons (Da) less than the theoretical values. This 4 Da discrepancy is likely due to the fact that two disulfide bonds are present within the molecule, which, if formed correctly, would result in four less hydrogen atoms. This result indicates that the full length βlg proteins have been successfully expressed and purified in an

intact, folded form. While it is not possible from the experimental mass values to determine if the correct disulfide bonds have formed, it is likely that the salt-fractionation step at low pH would have removed those that had formed incorrectly as they would likely not exhibit the same acid stability as the correctly folded product.

Table 3.1: The theoretical and calculated molecular masses of bovine and caprine βlg, utilising ExPASy ProtParam (Gasteiger *et al.*, 2005) and as assessed by mass spectrometry.

Protein	Theoretical molecular mass (Da)	Experimental molecular mass (Da)
Bovine βlg A	18498.4	18494.0
Bovine βlg B	18412.4	18407.8
Caprine βlg	18322.4	18318.4

3.2.4 Circular dichroism spectroscopy confirms that each βlg protein is folded

3.2.4.1 Circular dichroism spectroscopy

Circular dichroism is a phenomenon exhibited by chiral molecules that absorb left-handed and right-handed circularly-polarised light to different extents (Greenfield, 2007). This unequal absorption results in elliptically polarised light. A CD spectropolarimeter measures the difference in absorbance between the left-handed and right-handed circularly-polarised light and reports this in terms of ellipticity (θ) in millidegrees. When the ellipticity is reported as a function of wavelength we can obtain a CD spectrum.

Certain structural elements of proteins, such as α-helices and β-sheets, have characteristic far UV CD spectra (180 – 240 nm) (Greenfield, 2007). For instance, α-helices produce minima at 222 nm and 208 nm and a maximum at 193 nm, while β-sheets produce a minimum at 218 nm and a maximum at 195 nm. Disordered proteins, on the other hand, show very little ellipticity above 210 nm and a minimum at 195 nm, making far-UV CD spectroscopy a useful method for determining whether a protein is folded or not. Near-UV CD arises from the chirality of the

environment surrounding aromatic residues that absorb in the near-UV region (250 – 350 nm), namely Trp, Tyr and Phe. Differences between near-UV CD spectra of closely related proteins can suggest that these residues are in slightly different conformations, which can be used to infer structural differences.

3.2.4.2 Far-UV CD spectra of bovine and caprine βlg

Far-UV CD spectra were collected for bovine βlg A and B and caprine βlg to provide information on the arrangement of secondary structural elements within the proteins (Figure 3.5). The spectra were compared to the SMP180 reference set of proteins using the CONTIN algorithm (Provencher and Gloeckner, 1981; van Stokkum *et al*., 1990) within DichroWeb (Whitmore and Wallace, 2008) in order to estimate the secondary structure composition within each protein (Table 3.2).

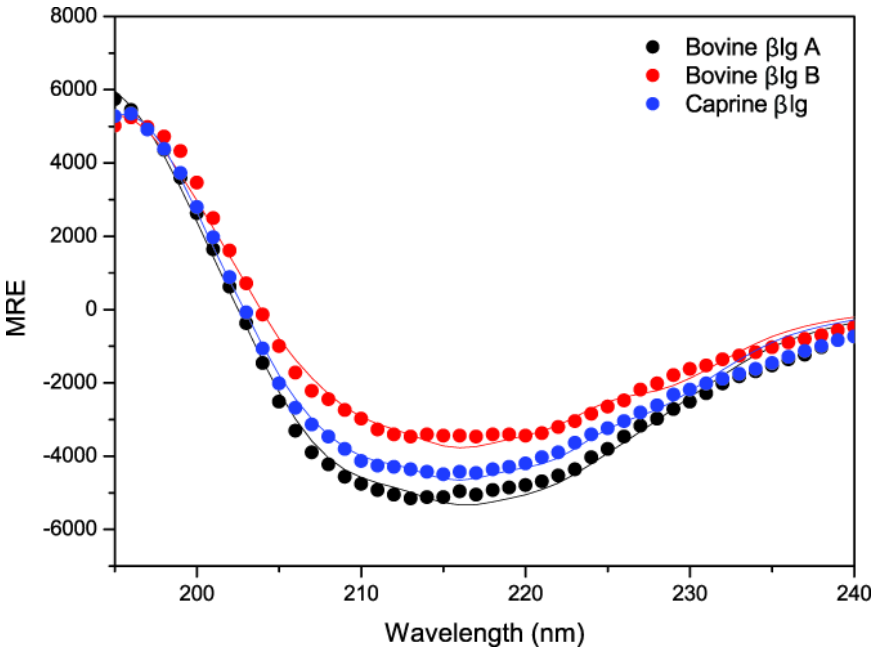


Figure 3.5: Far-UV CD spectra of bovine βlg A and B and caprine βlg at 0.05 mg/ml in water in a 2 mm cuvette. The spectra were collected between 180 and 260 nm but any data with a high tension voltage of over 600 V were excluded from analysis. The ellipticity data were converted to the mean residue ellipticity according to equations 2.1 and 2.2. Data (circles) are overlaid with a fit (line) against the SMP180 reference set of proteins using the CONTIN algorithm within DichroWeb.

Table 3.2: Secondary structure composition of bovine βlg A and B, and caprine βlg, estimated by comparing CD spectroscopy data to the SMP180 reference set of proteins using the CONTIN algorithm within DichroWeb.

Protein	Helices (%)	Sheets (%)	Turns (%)	Unordered (%)	rmsd
Bovine βlg A	12.2	36.2	12.1	39.6	0.068
Bovine βlg B	9.1	39.2	11.9	39.8	0.083
Caprine βlg	11	36.3	12.3	40.4	0.068

The spectra all show a dip at ~215 nm, reminiscent of a high proportion of β sheets within the structure (Greenfield, 2007). This is reflected in the secondary structure composition estimation, as shown in Table 3.2, which indicates that each βlg orthologue contains 36-39% β-sheet and 9-12% α-helix. These estimates are in close agreement with previous CD spectroscopic studies of bovine βlg that consistently report ~40% β-sheet and ~10% α-helix (Manderson *et al.*, 1999; Mehraban *et al.*, 2016; Qi *et al.*, 1997) and with crystal structures of bovine βlg (Figure 3.6) which show that the fold predominantly consists of β-sheets. While there is a 3% difference between bovine βlg A and B in terms of the percentages of helices and sheets estimated by CD spectroscopy, these differences are not observed in overlaid crystal structures of bovine βlg A and B (Figure 3.6). This variation is likely due to the uncertainty associated with secondary structure estimation from CD spectroscopy data. The differing intensities seen in the CD spectra between bovine A and B and caprine βlg are most likely due to differences in protein concentration, rather than significant differences in the secondary structural arrangements of these proteins.

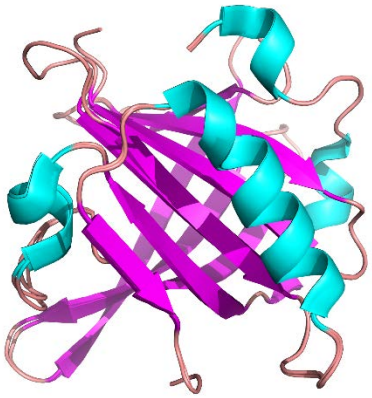


Figure 3.6: Crystal structure of overlaid monomeric bovine βlg A (PDB ID: 1QG5) and bovine βlg B (PDB ID: 1B8E) (Oliveira *et al.*, 2001). Cartoon representation with α-helices shown in blue, β-strands in magenta and loops in pink.

3.2.4.3 Near-UV CD spectra of bovine and caprine βlg

Near-UV CD spectra collected for bovine βlg A and B and caprine βlg strongly indicate that these recombinantly expressed proteins have each attained structures that are very similar to that of native bovine βlg. Features of a near-UV CD spectrum are due to the environment surrounding particular residues, such as Trp, Tyr and Phe. The two minima at 286 and 293 nm, seen here for each of the orthologues of βlg (Figure 3.7), are characteristic of native bovine βlg and have been seen in previous CD spectroscopic studies (Manderson *et al*., 1999; Townend *et al*., 1967). These troughs are attributed to residue tryptophan 19 (Manderson *et al*., 1999) and are very sensitive to the environment around this residue, which can therefore be used as a diagnostic feature to monitor the correct folding of these proteins.

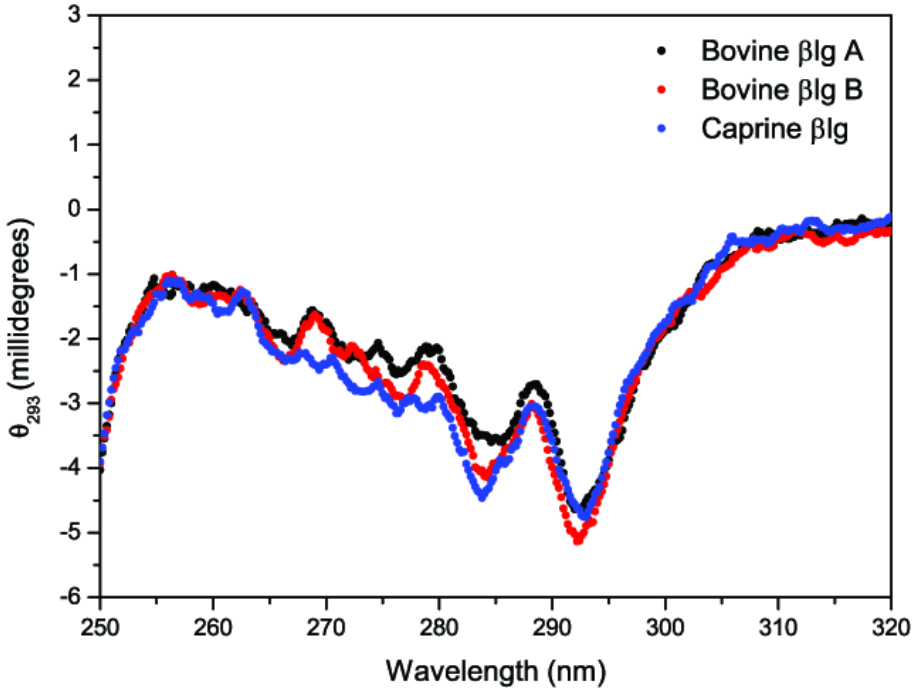


Figure 3.7: Near-UV CD spectra of bovine A and B βlg and caprine βlg at 1 mg/ml in 0.02 M sodium phosphate, 0.1 M sodium chloride, pH 6.7 buffer in a 1 cm cuvette. The spectra were collected between 250 and 320 nm.

3.3 Summary

Through the utilisation of a co-expressed disulfide bond isomerase and the environment within *E. coli* Origami (DE3) cells, caprine βlg has been successfully recombinantly expressed for the first time. The same purification steps that have been successful in purifying bovine βlg variants are also effective in purifying caprine βlg to homogeneity. The exploitation of the acid stability of these proteins has an added benefit in that it not only removes contaminating proteins, but that it also likely removes any βlg that has not attained the correct conformation.

The experimentally determined molecular masses of these proteins indicate that they have been expressed as full length constructs, with the correct number of disulfide bonds. The close agreement of both the far-UV and near-UV CD spectra between orthologues suggests that caprine βlg shares a very similar architecture with that of its closely related bovine orthologues. It also supports that recombinant protein expression is appropriate in this case for producing

correctly folded proteins that serve as suitable models for these proteins produced naturally in milk.

3.4 References

Ariyaratne, K.A.N.., Brown, R., Dasgupta, A., de Jonge, J., Jameson, G.B., Loo, T.S., Weinberg, C., Norris, G.E., 2002. Expression of bovine β-lactoglobulin as a fusion protein in *Escherichia coli*: a tool for investigating how structure affects function. International Dairy Journal 12, 311–318.

Batt, C.A., Rabson, L.D., Wong, D.W., Kinsella, J.E., 1990. Expression of recombinant bovine B-lactoglobulin in *Escherichia coli*. Agricultural and Biological Chemistry 54, 949–955.

Bessette, P., Åslund, F., Beckwith, J., Georgiou, G., 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proceedings of the National Academy of Sciences 96, 13703–13708.

Chatel, J.M., Bernard, H., Clement, G., Frobert, Y., Batt, C.A., Gavalchin, J., Peltres, G., Wal, J.M., 1996. Expression, purifcation and immunochemical characterization of recombinant bovine beta-lactoglobulin, a major cow milk allergen. Molecular Immunology 33, 1113–1118.

Cho, Y., Batt, C.A., Sawyer, L., 1994. Probing the retinol-binding site of bovine beta-lactoglobulin. Journal of Biological Chemistry 269, 11102–11107.

Denton, H., Smith, M., Husi, H., Uhrin, D., Barlow, P., 1998. Isotopically Labeled Bovine β-Lactoglobulin for NMR Studies Expressed inPichia pastoris. Protein Expression and Purification 14, 97–103.

Fogliano, V., Monti, S.M., Visconti, A., Randazzo, G., Facchiano, A.M., Colonna, G., Ritieni, A., 1998. Identification of a beta-lactoglobulin lactosylation site. Biochim. Biophys. Acta 1388, 295–304.

Gasteiger, E., Hoogland, C., Gattiker, A., S, Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server. Humana Press Inc., Totawa, NJ.

Greenfield, N., 2007. Using circular dichroism spectra to estimate protein secondary structure. Nature Protocols 1, 2876–2890.

Hyttinnen, J., Korhonen, V.P., Hitunen, M.O., Myohanen, S., Jainne, J., 1998. High level expression of the beta-lactoglobulin gene in transgenic mice. Journal of Biotechnology 61, 191–198.

Kim, T.R., Goto, Y., Hirota, N., Kuwata, K., Denton, H., Wu, S.Y., Sawyer, L., Batt, C.A., 1997. High-level expression of bovine beta-lactoglobulin in Pichia pastoris and characterization of its physical properties. Protein Engineering 10, 1339–1345.

Kurokawa, Y., Yanagi, H., Yura, T., 2000. Overexpression of protein disulfide isomerase DsbC stabilizes multiple-disulfide-bonded recombinant protein produced and transported to the periplasm in *Escherichia coli*. Applied and environmental microbiology 66, 3960–5.

Manderson, G., Creamer, L., Hardman, M., 1999. Effect of heat treatment on the circular dichroism spectra of bovine beta-lactoglobulin A, B, and C. Journal of Agricultural and Food Chemistry 47, 4557–67.

Mehraban, M.H., Odooli, S., Yousefi, R., Roghanian, R., Motovali-Bashi, M., Moosavi-Movahedi, A.-A.A., Ghasemi, Y., 2016. The interaction of beta-lactoglobulin with ciprofloxacin and kanamycin; a spectroscopic and molecular modeling approach. Journal of biomolecular structure & dynamics 1–11.

Mizumachi, K., Kurisaki, J., Tsuji, N.M., 1993. High level expression of recombinant bovine beta-lactoglobulin in insect cells. Netherlands, pp. 115–120.

Oliveira, K.M., Valente-Mesquita, V.L., Botelho, M.M., Sawyer, L., Ferreira, S.T., Polikarpov, I., 2001. Crystal structures of bovine β-lactoglobulin in the orthorhombic space group C2221. European Journal of Biochemistry 268, 477–484.

Ponniah, K., Loo, T., Edwards, P., Pascal, S., Jameson, G., Norris, G., 2010. The production of soluble and correctly folded recombinant bovine β-lactoglobulin variants A and B in *Escherichia coli* for NMR studies. Protein Expression and Purification 70, 283–289.

Provencher, S., Gloeckner, J., 1981. Estimation of globular protein secondary structure from circular dichroism. Biochemistry-us 20, 33–37.

Qi, X.L., Holt, C., McNulty, D., Clarke, D.T., Brownlow, S., Jones, G.R., 1997. Effect of temperature on the secondary structure of beta-lactoglobulin at pH 6.7, as determined by CD and IR spectroscopy: a test of the molten globule hypothesis. The Biochemical journal 324, 341–6.

Townend, R., Kumosinski, T., Timasheff, S., 1967. The circular dichroism of variants of β-lactoglobulin. The Journal of Biological Chemistry 242, 4538–4545.

Van Stokkum, I.H., Spoelder, H.J., Bloemendal, M., van Grondelle, R., Groen, F.C., 1990. Estimation of protein secondary structure and error analysis from circular dichroism spectra. Analytical Biochemistry 191, 110–118.

Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 89, 392–400.

Chapter Four

4. A Structural Characterisation of Caprine β-lactoglobulin

Part of the work presented in this chapter has been published in Crowther, J., Lasse, M., Suzuki, H., Kessans, S., Loo, T., Norris, G., Hodgkinson, A., Jameson, G., Dobson, R. (2014) Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin. FEBS Letters, 588, 3816-3822.

4.1 Introduction

4.1.1 Bovine βlg contributes to cow milk allergies

Allergies are caused when the immune system produces immunoglobulin E (IgE) antibodies against otherwise harmless proteins. The structure of bovine βlg in complex with an IgE antibody fragment has been solved (Figure 4.1) (Niemi *et al*., 2007), which supports claims that bovine βlg is capable of contributing to cow milk allergies (Kapila *et al*., 2013). While both cow milk and goat milk are capable of eliciting allergic reactions in sensitised patients, five times more goat milk is required to induce the same response (Bellioni-Businco *et al*., 1999).

There are several reasons why goat milk may have a reduced allergenic burden compared to cow milk. Studies have shown that during *in vitro* digestion caprine βlg is digested more efficiently than bovine βlg (Almaas *et al*., 2006). It may be the case, therefore, that in goat milk a decreased amount of allergen passes through the digestive tract in its native allergenic

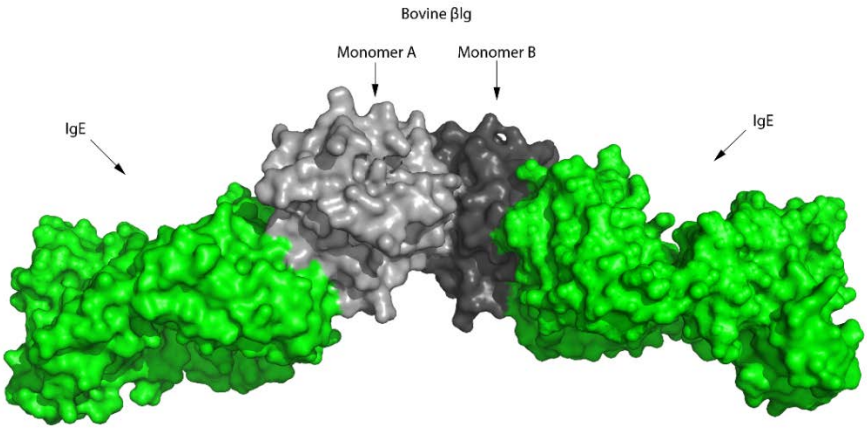


Figure 4.1: Bovine βlg (grey) in complex with two IgE antibodies (green) (PDB ID: 2R56 (Niemi *et al.*, 2007)).

state. What is unknown is whether the enhanced digestibility of goat milk proteins compared to cow milk proteins is due to species-specific sequence substitutions that reduce protein stability, or whether compositional differences between the milks are responsible for the different extents of protein digestion seen.

It may be that the differences in primary structure between cow and goat milk protein allergens alter the conformation of these proteins enough to affect the affinity of IgE antibodies. While a wealth of structural knowledge exists for bovine βlg, far fewer studies have focused on the caprine orthologue of βlg and at the commencement of this study no structures had been solved for caprine βlg.

It is also important to consider the role that the quaternary state plays in the immunogenicity of allergens. Allergens need to bind to at least two IgE antibodies in order to elicit an allergic response, as this allows the cross-linking of two IgE-receptor complexes, which is required to initiate the allergic reaction (Niemi *et al*., 2007). In the case of bovine βlg there are two epitope sites per dimer. This means only one type of IgE antibody is required for cross-linking, which increases the potency of bovine βlg to act as an allergen when it is in its dimeric form. The behaviour of caprine βlg in solution is thus far unknown.

4.1.2 Quaternary structure of bovine βlg

The oligomerisation of bovine βlg has been thoroughly investigated using a range of techniques including analytical ultracentrifugation, isothermal titration calorimetry and small-angle X-ray scattering (Bello *et al*., 2011; McKenzie and Sawyer, 1972; Mercadante *et al*., 2011; Sakurai and Goto, 2002; Sakurai *et al*., 2001). The results of these studies have been comprehensively reviewed by Mercadante *et al*. (2011), but in general, bovine βlg exists in a monomer-dimer equilibrium. This equilibrium can be altered by pH and by ionic strength. At the pH and protein concentrations found in milk, bovine βlg is likely to be a dimer.

The dimer interface of bovine βlg, shown in Figure 4.2 is composed of an intermolecular β-sheet formed by the anti-parallel association of the 'I' β-strands of each monomer, along with electrostatic interactions and hydrogen bonding between Asp 33 and Arg 40, located on the AB loops. Sakurai and Goto (2002) investigated the importance of these elements for dimer formation using analytical ultracentrifugation. They engineered bovine βlg mutants that either interfered with the intermolecular β-sheet or the interactions between the AB loops. All of these mutants led to dissociation of the dimer, indicating the essentiality of these elements to stabilisation of the dimer.

At low pH the secondary and tertiary structure of bovine βlg are well maintained, while the quaternary structure is very sensitive to ionic strength (Mercadante *et al*., 2011; Sakurai and Goto, 2002). At pH values less than 3 and at low salt concentrations (< 10 mM) bovine βlg is predominantly monomeric. This is due to the positive charge on the monomers at low pH, which forces the subunits to separate unless there are enough ionic species present to screen and neutralise the charges near the dimer interface (Mercadante *et al*., 2011).

Some orthologues, such as equine βlg, exist as monomers at neutral pH despite containing the same 'I' strands and AB loop residues as bovine βlg (Ohtomo *et al*., 2012). From a series of studies aimed at interconverting between the monomeric and dimeric forms of βlg (Kobayashi *et al*., 2002; Ohtomo *et al*., 2012, 2011; Sakurai, 2002), it was deduced that the entire arrangement of the secondary structural elements and loops of βlg, in addition to hydrophobic interactions, hydrogen bonds between I-strands, and electrostatic interactions and hydrogen bonds between AB loops, are all necessary for dimerisation. Further, it is tempting to speculate that protein dynamics also play a role in dimer formation.

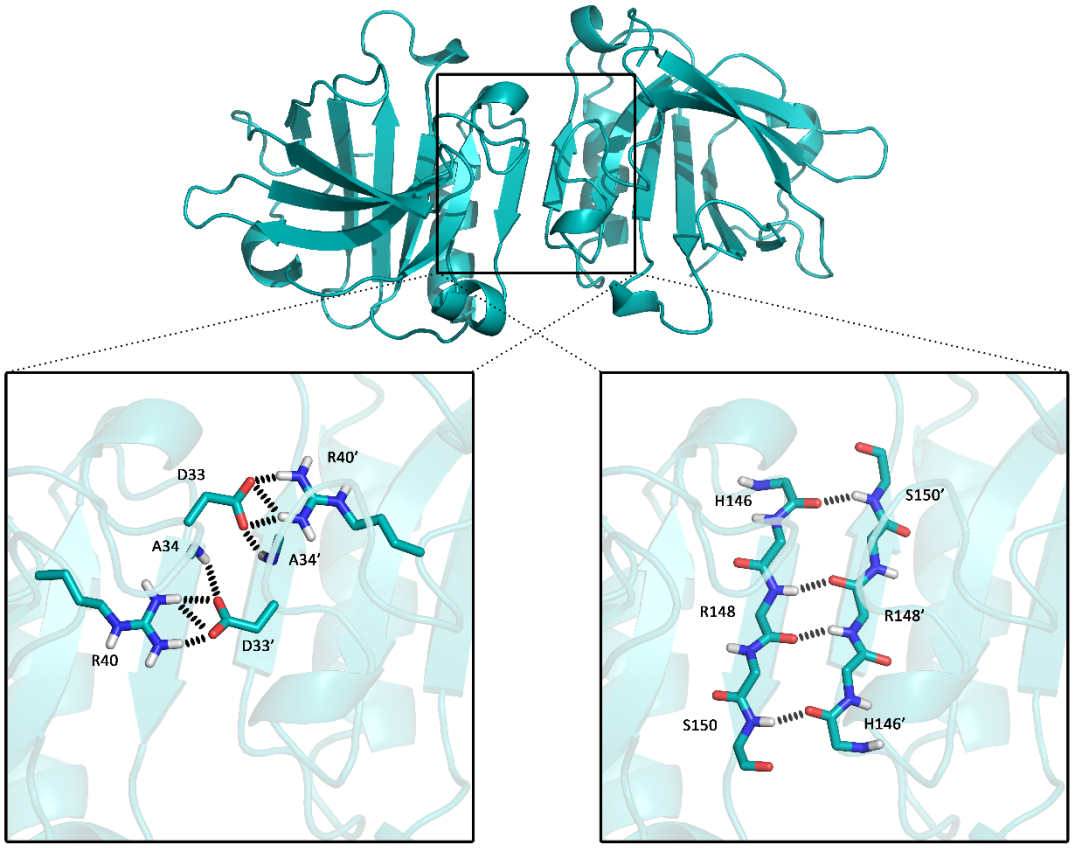


Figure 4.2: Dimer interface of bovine βlg (PDB ID: 1BEB (Brownlow *et al.*, 1997)). Close-up views of AB-loops and I-strands are shown where inter subunit hydrogen bonds and electrostatic interactions between side chains and main chains can be seen. (Crowther *et al.*, 2016) © under CC BY 3.0 license.

4.1.2.1 Quaternary state characterisation of bovine βlg by analytical ultracentrifugation

Mercadante *et al*. (2011) carried out a detailed characterisation of the quaternary-state association behaviour of bovine βlg A and B by analytical ultracentrifugation. They examined a pH range of 2.5 to 7.5 with a constant ionic strength of 100 mM NaCl. At each pH the protein was analysed at three concentrations (5, 15 and 45 µM). The sedimentation velocity analyses were fitted to continuous sedimentation coefficient distributions.

Interestingly, at pH 2.5, 3.5, 6.5 and 7.5 the weight-averaged sedimentation coefficient increases with increasing protein concentration, indicating a monomer-dimer equilibrium. However, at pH 4.5 and 5.5, the weight-averaged sedimentation coefficient stays the same, consistent with a dimer prevailing at each concentration used in the experiment. This suggests

that the dimer of bovine βlg is more strongly associated at pH 4.5 and 5.5, values that lie near the isoelectric point of the protein, than at pH 2.5, 3.5, 6.5 and 7.5 (Mercadante *et al*., 2011).

Of the eight amino acid substitutions between bovine βlg A and caprine βlg (L1I, D53N, D64G, V118A, D130K, S150A, E158G and I162V) four involve the substitution of a negatively charged residue (at neutral pH) for a neutral residue (D53N, D64G, E158G) or a positively charged residue (D130K). This results in a greater number of positively charged amino acids and a lesser number of negatively charged amino acids in caprine βlg compared to the bovine A and B βlg variants (Table 4.1). This results in caprine βlg carrying a lesser negative charge at pH 7 and having a higher isoelectric point (i.e. the pH at which the protein carries a net neutral charge) than the bovine orthologues. These substitutions have the potential to affect the self-association properties of caprine βlg across a pH range.

Table 4.1: Theoretical isoelectric point and net charge values for bovine and caprine βlg.

	Bovine βlg A	Bovine βlg B	Caprine βlg
Theoretical pI (ExPasy)	4.76	4.83	5.29
Theoretical pI (Sednterp)	5.01	5.08	5.51
Estimated charge at pH 7 (Sednterp)	-8.95	-7.95	-3.96
Number of positively charged amino acids	18	18	19
Number of negatively charged amino acids	27	26	23

4.1.3 Overview

It is possible that the amino acid sequence differences present between bovine and caprine βlg alter the conformations of these proteins such that they are recognised to a different extent by IgE antibodies. Alternatively, the quaternary state association behaviour of caprine βlg in solution may be dissimilar to that of bovine βlg. The aim of this chapter was to elucidate the structure of the caprine orthologue of βlg and to examine its behaviour in solution. This information is necessary to understand how these proteins are presented to the immune system during digestion.

This chapter details the characterisation of the solution-state behaviour of bovine and caprine βlg. While a monomer-dimer equilibrium is observed between pH values of 3.5 and 7.5, caprine βlg is predominantly monomeric at pH 2.5. It is thought that the increased positive charge on each monomer of caprine βlg at this pH compared to bovine βlg A and B may shift the equilibrium towards the monomeric state. This has the potential to affect how these proteins are presented to the immune system during digestion.

A structural characterisation of caprine βlg was undertaken in order to understand whether there are any significant structural differences between the bovine and caprine orthologues of βlg, due to variations in the amino acid sequences of these proteins. Collection of a complete ultra-high resolution X-ray diffraction data set has allowed the three-dimensional structure of caprine βlg to be solved for the first time. This reveals that the caprine orthologue of βlg adopts a similar fold to that of bovine βlg. Small-angle X-ray scattering data reinforces that the crystal structure is an accurate model of this protein structure in solution. Coupled with this structural knowledge, differences in the behaviour of caprine βlg in solution can be rationalised in terms of the sequence differences between the bovine and caprine orthologues.

4.2 Results and discussion

4.2.1 Quaternary state characterisation of caprine βlg by analytical centrifugation

4.2.1.1 Analytical ultracentrifugation

Analytical ultracentrifugation is a versatile technique for the analysis of molecules in solution. The instrument is equipped with transparent centrifuge cells and various optics systems which allow direct monitoring of the sedimentation of molecules during centrifugation. This provides information about the shape and the size of these molecules which can be useful, for example, in the quaternary state characterisation of proteins (Lebowitz *et al*., 2002).

In sedimentation velocity analyses of solutes the fundamental measurement is the concentration of the solute as a function of the radial position (Laue and Stafford, 1999). As a centrifugal force is applied to the solution the concentration of solute becomes depleted at the top of the cell and forms a boundary that moves toward the bottom of the cell (Lebowitz *et al*., 2002). The rate at which the solute sediments is dependent upon the size and shape of the molecule. This is described by the sedimentation coefficient, s, which is given by the Svedberg equation:

Equation 4.1:

$$s = \frac{u}{\omega^2 r} = \frac{M(1 - \bar{v}\rho)}{N_A f}$$

Where u is the observed radial velocity of the molecule, $\omega^2 r$ is the centrifugal field, M is the molar mass of the molecule, $\bar{v}$ is the partial specific volume, ρ is the density of the solvent, N_A is Avogadro's number and f is the frictional coefficient.

Diffusion toward the top of the cell (i.e. down the concentration gradient that forms during sedimentation) causes the concentration boundary to spread, thus what is actually observed is a concentration distribution over the length of the cell. Concentration distribution profiles taken at sequential time points during centrifugation allow direct observation of the movement of the concentration boundary with time. The Lamm equation is a function that describes the concentration distribution by taking into account sedimentation and diffusion. Software, such as SEDFIT, can be used to model sedimentation data by solving the Lamm

equation for each profile and fitting the data to a continuous sedimentation coefficient distribution.

4.2.1.2 Caprine βlg displays a dependence of quaternary state on pH that differs from that of bovine βlg.

A sedimentation velocity analysis replicating the conditions of that conducted by Mercadante *et al*. (2011) (see section 4.1.2.1 for details) was carried out in order to compare the solution behaviour of caprine βlg to that of the bovine orthologues of βlg. Caprine βlg was analysed at three concentrations (5, 15 and 45 μM) at pH 2.5, 3.5, 4.5, 5.5, 6.5, and 7.5. For each sample the sedimentation velocity data was fitted to a continuous sedimentation coefficient distribution (Figure 4.3) and a continuous mass distribution (Figure 4.4). The weight-averaged sedimentation coefficients were obtained by integrating between s values of 1 and 5 (Table 4.2).

The sedimentation velocity analysis of caprine βlg over a wide pH range shows some interesting differences when compared to bovine βlg under the same experimental conditions. Mercadante *et al*. report that at pH 2.5, bovine βlg is in a monomer-dimer self-association (over a concentration range of 5 – 33 μM). In contrast, caprine βlg is seen here as a predominantly monomeric species at this pH (over an even broader concentration range of 7 – 58 μM).

In order to ascertain that the differences seen at pH 2.5 were not due to experimental set-up, samples of bovine βlg A and caprine βlg were analysed concurrently at pH 2.5 in the same buffer. This ensured that all experimental conditions were identical across samples. Analysis of the sedimentation velocity data confirms that bovine βlg A does indeed exhibit different quaternary-state association behaviour than caprine βlg at low pH (Figure 4.5). This also serves to show that the experimental set-up used here is a good replicate of the experimental set-up utilised by Mercadante *et al*. (2011), which allows for direct comparisons to be made between these studies. While the weight-averaged s value of bovine βlg A increases from 2.03 to 2.30 (from 5.8 μM to 52.8 μM), the weight-averaged s value of caprine βlg increases only slightly from 1.88 to 1.98 (from 5.43 μM to 51.5 μM). When fitted to a *c*(*M*) distribution, the mass of the caprine βlg species at the highest concentration is estimated to be 17,800 Da, which is in

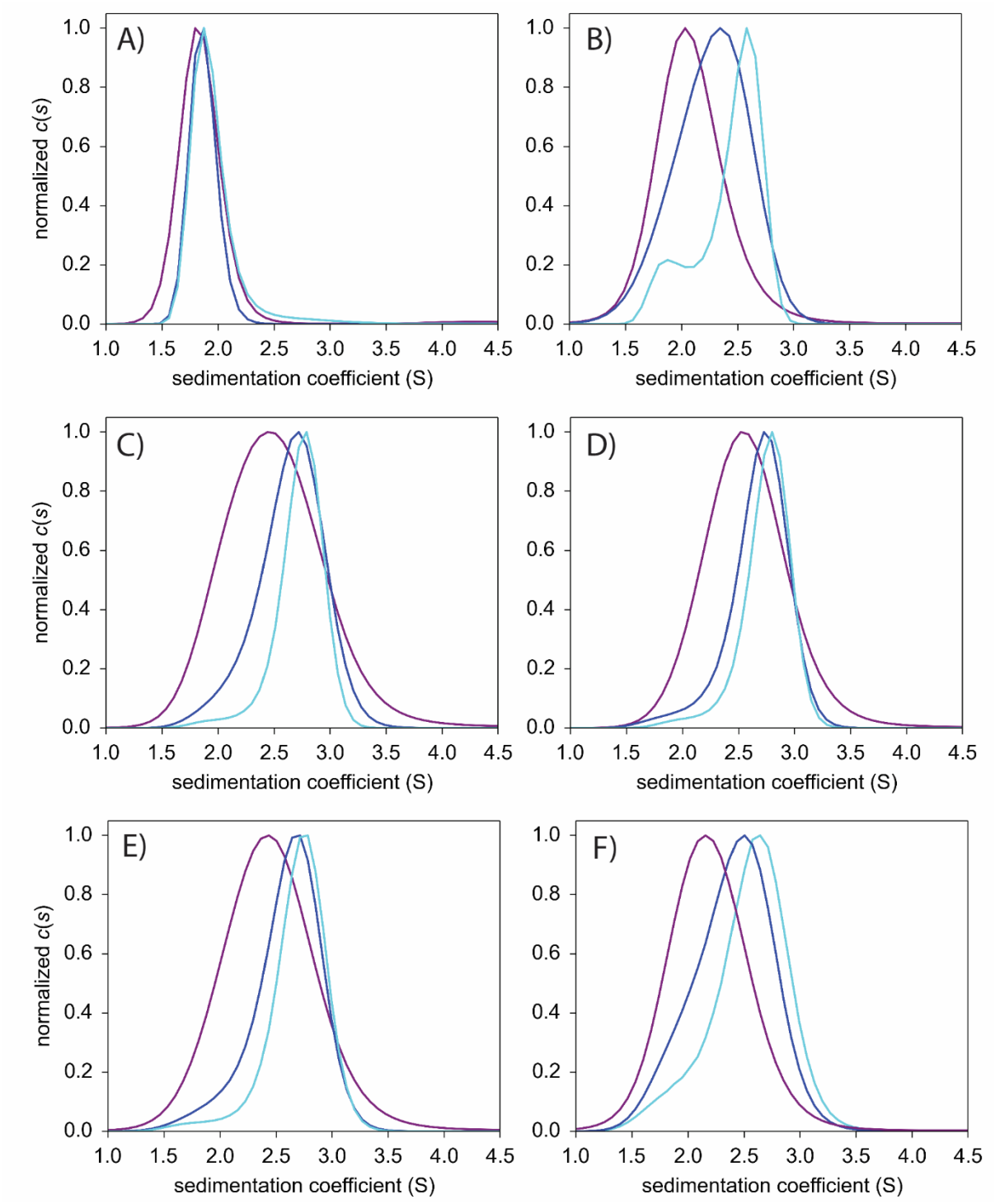


Figure 4.3: Normalised c($s_{20,w}$) distribution plots for caprine βlg over the pH range 2.5 – 7.5. The three protein concentrations tested were ~5 µM (purple), ~15 µM (blue) and ~45 µM (cyan); actual concentrations and fit statistics are shown in Table 4.2. A) pH 2.5, B) 3.5, C) 4.5, D) 5.5, E) 6.5, F) 7.5. Two buffers were used: 20 mM sodium citrate, 100 mM NaCl (pH 2.5 – 5.5) or 20 mM MOPS, 100 mM NaCl (pH 6.5 – 7.5).

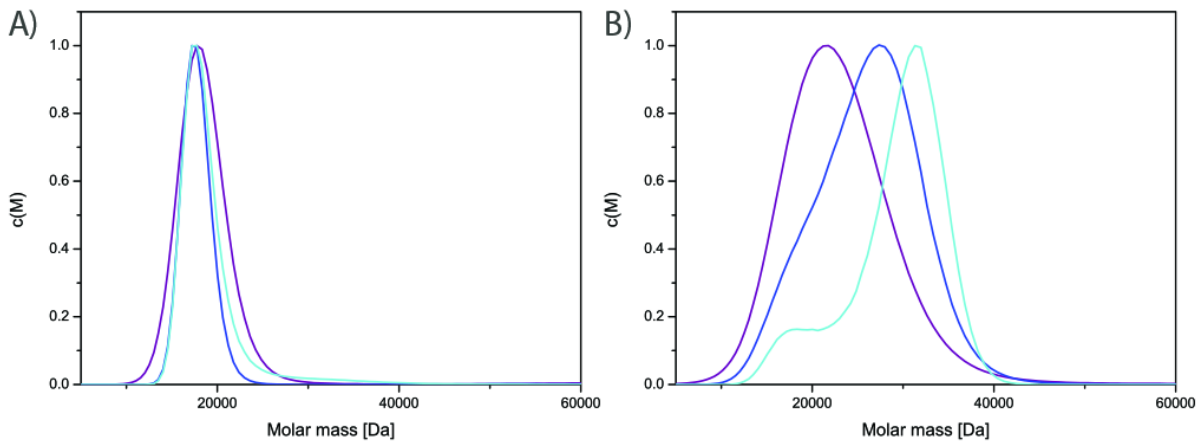


Figure 4.4: Continuous mass distribution plots for caprine βlg at A) pH 2.5 and B) pH 7.5. See Table 4.2 for fit statistics.

closer agreement with the monomeric mass (18,318 Da determined experimentally, see section 3.2.3) than with the dimeric mass (36,636 Da). This suggests that the $K_D^{2\text{-}1}$ for caprine βlg at this pH is above the range of concentrations investigated (>58 µM).

The behaviour of caprine βlg at pH 4.5 and 5.5 also differs from that seen for bovine βlg. For bovine βlg, a concentration dependent change in the oligomeric state is observed at pH values of 2.5, 3.5, 6.5 and 7.5. At pH 4.5 and 5.5, however, the weight averaged s does not increase with concentration. When fitted to a *c*(*M*) distribution, the mass fits closely with the mass of the dimer. This suggests that the dimer is the predominant species at these pH values, and that the $K_D^{2\text{-}1}$ is much lower than the concentrations investigated (< 5 µM) (Mercadante *et al.*, 2011). This is in contrast to what is seen for caprine βlg; at pH values of 3.5, 4.5, 5.5, 6.5 and 7.5, the weight averaged s increases with protein concentration. The asymmetric peaks at lower concentrations tail to higher s values, while at higher concentrations they tail to smaller s values. This is indicative of a monomer-dimer self-association, and of $K_D^{2\text{-}1}$ values within the range of concentrations examined here. At pH values of 4.5, 5.5 and 6.5, the peaks for caprine βlg are shifted slightly further to the right than at pH 3.5 and 7.5. This suggests that the $K_D^{2\text{-}1}$ values are slightly lower at these pH values, but not as low as those for bovine βlg.

Table 4.2: Sedimentation velocity parameters of caprine βlg.

Sample	Model	Weight-averaged $s_{20,w}$ (S)	Weight-averaged Mass (kDa)	f/f0	Runs-Z test score	rmsd
Citrate pH 2.5						
6.7 µM	*c*(*s*)	1.99		1.24	0.75	0.005
20.1 µM	*c*(*s*)	1.91		1.19	0.85	0.006
58.1 µM	*c*(*s*)	2.04		1.19	4.84	0.009
58.1 µM	*c*(*M*)		17.8	1.15	3.96	0.009
Citrate pH 3.5						
7.1 µM	*c*(*s*)	2.28		1.13	0.77	0.005
21.8 µM	*c*(*s*)	2.29		1.14	0.26	0.006
57.2 µM	*c*(*s*)	2.42		1.15	8.29	0.009
MOPS pH 4.5						
6.2 µM	*c*(*s*)	2.59		1.20	0.24	0.007
18.5 µM	*c*(*s*)	2.66		1.12	1.08	0.007
57.2 µM	*c*(*s*)	2.72		1.14	3.56	0.013
MOPS pH 5.5						
5.6 µM	*c*(*s*)	2.64		1.20	1.51	0.007
17.2 µM	*c*(*s*)	2.68		1.16	0.61	0.007
54.2 µM	*c*(*s*)	2.74		1.20	9.00	0.011
MOPS pH 6.5						
5.7 µM	*c*(*s*)	2.54		1.17	0.92	0.007
17.7 µM	*c*(*s*)	2.60		1.15	0.65	0.007
53.3 µM	*c*(*s*)	2.71		1.20	2.73	0.014
MOPS pH 7.5						
6.4 µM	*c*(*s*)	2.29		1.20	0.52	0.007
17.4 µM	*c*(*s*)	2.41		1.20	0.62	0.007
54.7 µM	*c*(*s*)	2.57		1.28	4.69	0.014
54.7 µM	*c*(*M*)		29.6	1.24	5.13	0.013

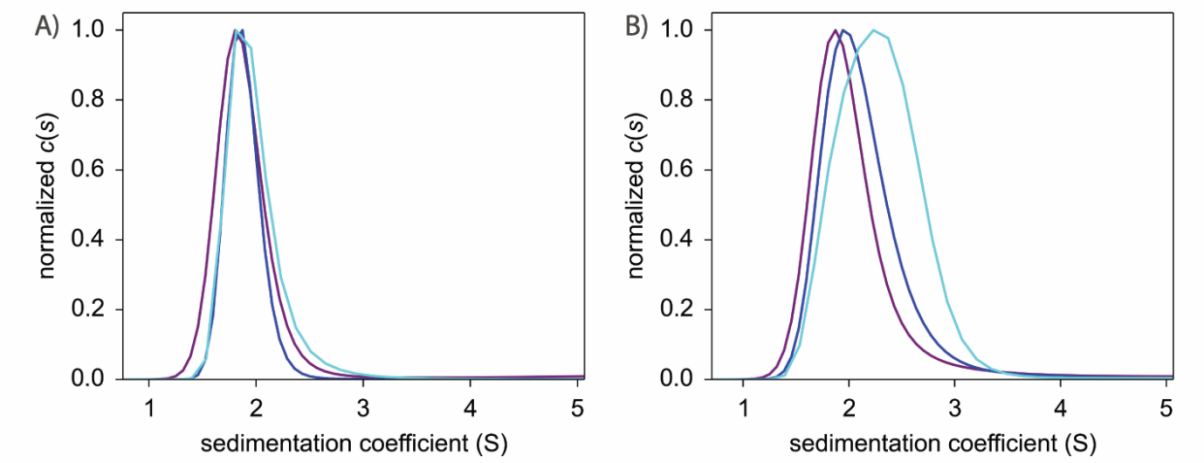


Figure 4.5: Normalised c($s_{20,w}$) distribution plot of A) caprine βlg at pH 2.5 (5.43 µM, 17.3 µM and 51.5 µM) and B) bovine βlg A at pH 2.5 (5.77 µM, 17.4 µM and 52.8 µM). The rmsd for each plot are as follows: 0.009, 0.009, 0.011, 0.009, 0.009 and 0.012. The runs test-z scores are as follows: 0.44, 0.86, 1.64, 0.18, 0.11 and 1.13.

Both caprine and bovine βlg appear to exist predominantly as dimers in solution at 50 µM and pH 6.5 (Figure 4.3, (Mercadante *et al*., 2011)). This suggests that at the pH and concentration likely to be encountered in milk, both orthologues would likely be in the dimeric form (pH 6.6-6.7 and 3 mg/mL or ~150 µM (Kontopidis *et al*., 2004; le Maux *et al*., 2014). In contrast, at the low pH values potentially encountered during digestion, it is possible that these orthologues may experience a different oligomeric state. However, association at low pH is also highly dependent on ionic strength.

4.2.2 Atomic structure of caprine βlg

4.2.2.1 Crystallisation of caprine βlg

Elucidation of an ultra-high resolution crystal structure of caprine βlg has revealed that caprine βlg adopts a fold very similar to that of the bovine orthologues of bovine βlg. The Morpheus crystallisation screen (Molecular Dimensions Ltd.) was successful in producing an extremely high quality crystal of caprine βlg (Figure 4.6). The condition in which the crystal formed (condition D1: 0.2 M 1,6-Hexanediol, 0.2 M 1-Butanol, 0.2 M 1,2-Propanediol, 0.2 M 2-Propanol, 0.2 M 1,4-Butanediol, 0.2 M 1,3-Propanediol, 0.1 M imidazole, 0.1 M MES, 20% v/v

PEG 500 MME and 10% w/v PEG 20000) contained a mixture of aliphatic alcohols at a low salt concentration, in contrast to the very high salt conditions under which bovine βlg is traditionally crystallised. The crystal formed after two weeks at room temperature (~20°C). The pH of crystallisation was estimated from combining equal volumes of 50 mM Tris at pH 7.5 (protein solution buffer) with 100 mM MES/imidazole at pH 6.5 (reservoir solution buffer) to be around 6.8 – 6.9.

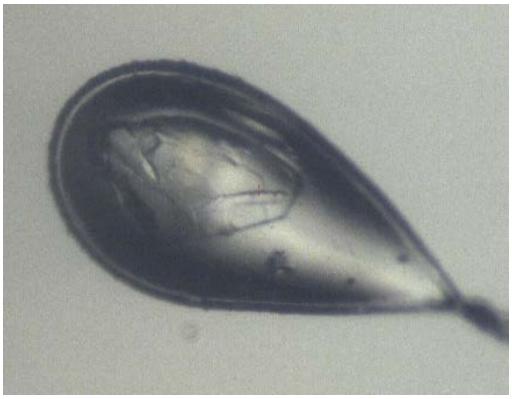


Figure 4.6: Crystal of caprine βlg mounted in a litho-loop.

4.2.2.2 X-ray diffraction data collection

For data collection the crystal was harvested in a litho-loop and flash-cooled in liquid nitrogen. The presence of PEG 20,000 and PEG MME 550 in the crystallisation condition removed the need to add any other cryo-protectant. The crystal was mounted onto the beamline goniometer in a cold nitrogen stream and rotated while monochromatic collimated X-rays were applied to it. This produced a diffraction pattern of regularly spaced reflections corresponding to the ordered array of atoms within the crystal (Figure 4.7). Diffraction data was collected in 1° steps over a 360° pass allowing collection of a complete data set to a very high resolution of 1.09 Å.

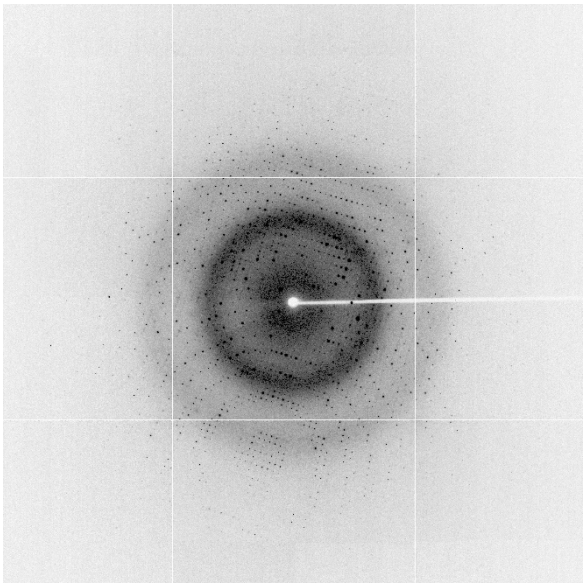


Figure 4.7: Diffraction pattern from a crystal of caprine βlg

4.2.2.3 Data processing

During data processing, the reflections were indexed, integrated and scaled. Indexing provides information about the dimensions of the unit cell and the symmetry of the system, while integration and scaling combines all of the information from the 360 images and normalises their intensities (Rhodes, 2006). In this case, the crystal belonged to space group P21. The asymmetric unit contained two monomers of caprine βlg based on a Matthew's coefficient of 2.14 Å-3 Da-1 and a solvent content of 42%. Data to 1.17 Å were used, as this corresponded to a mean ratio of I/σ(I) greater than 2 in the outer shell of reflections. Data collection and processing statistics are summarised in Table 4.3

4.2.2.4 Molecular replacement

During data collection, the amplitude and direction of the diffracted X-rays are recorded, but their phase information is lost. This information is required to complete the Fourier transformation to model the positions of atoms in the original crystal. The phase problem can

Table 4.3: X-ray diffraction data-collection, processing and refinement statistics of caprine βlg. Values in brackets pertain to the highest resolution shell.

	Caprine βlg
Data collection	
Space group	$P2_1$
Wavelength (Å)	0.9537
Number of images	360
Oscillation range per image (°)	1.0
Detector	ADSC Quantum 315r
Cell dimensions	
a, b, c (Å)	52.3, 54.6, 56.4
α, β, γ (°)	90, 107.3, 90
Resolution (Å)	23.30–1.17 (1.23–1.17)
R_{sym} †	0.081 (0.694)
R_{meas} §	0.087 (0.747)
R_{pim} ‡	0.032 (0.276)
I/σI	12.2 (2.6)
Total observations	729,152 (104,869)
Unique reflections	101,724 (14,688)
Completeness (%)	99.6 (99.1)
Redundancy	7.2 (7.1)
Wilson B-factor (Å^2)	10.3
Matthews Coefficient (Å^3 Da^{-1})	2.14
Solvent content (%) (2 mol. per asu)	42
Refinement	
Resolution (Å)	23.30–1.17 (1.23–1.17)
Reflections used in refinement	99,665
R_{free} reflections	2,042
R_{work}	0.132 (0.242)
R_{free}	0.159 (0.291)
Protein molecules in asymmetric unit	2
Protein residues	324
Total atoms	
Protein	1729
Ligand/ion	6
Water	478
Mean anisotropic B-factor (Å^2)	
Protein	14.2
Ligand/ion	28.8
Water	29.7
r.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.655

† $R_{sym} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$

§ $R_{meas} = \sum_{hkl} [N/(N-1)]^{½} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$

‡ $R_{pim} = \sum_{hkl} [1/(N-1)]^{½} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$

be solved using a technique called molecular replacement, which uses the phases from structure factors of a known, similar, protein structure (Rhodes, 2006). The position and orientation of the phasing model must be arranged in the new unit cell such that it superimposes on that of the new structure. Structure factors (which include amplitude as well as phase information) can be calculated for the phasing model and the phases can be used as the initial estimates of the phases of the new structure (Rhodes, 2006).

Initial phases were estimated by molecular replacement using the programme MOLREP and the crystal structure of bovine βlg A (PDB ID: 1BSO (Qin *et al*., 1998), 95% sequence identity) as the search model. The initial round of refinement following molecular replacement resulted in a decrease of the R_{work} and R_{free} values from 0.4594 to 0.2812 and 0.4550 to 0.2930, respectively. This suggests that the orientation and number of molecules in the unit cell, as determined by molecular replacement, were correct.

4.2.2.5 Model building

Having solved the crystal structure, iterative rounds of model building were conducted, during which the protein chains of caprine βlg were built into the clouds of electron density. Due to the high resolution of the data set the model was refined anisotropically. A final model was created with an R_{free} of 0.159 and R_{work} of 0.132. The model has good geometry and has only two outliers in the Ramachandran plot: Tyr99, which is a highly conserved residue in βlg orthologues (see Figure 1.1) and forms part of a gamma turn, and Ala34, which forms part of the dimer interface and is held in an unfavourable conformation by two hydrogen bonds, as seen in bovine βlg structures (Brownlow *et al*., 1997).

4.2.2.6 Structural features of caprine βlg

The atomic-resolution structure of caprine βlg reveals that it adopts a lipocalin fold similar to that of the closely related bovine orthologues. As shown in a three-dimensional cartoon form in Figure 4.8, each subunit consists predominantly of an antiparallel β-sheet formed by eight β-strands, A to H, wrapped around to form a flattened calyx. This cavity, in βlg structures from other species, has been seen to bind small hydrophobic molecules (Qin *et al*., 1998). No ligands

were observed to be bound in the calyx in this structure. The calyx is flanked on its outer surface by a three-turn α-helix. The ninth β-strand, I, and the AB loop form the dimer interface, as seen in structures of bovine βlg.

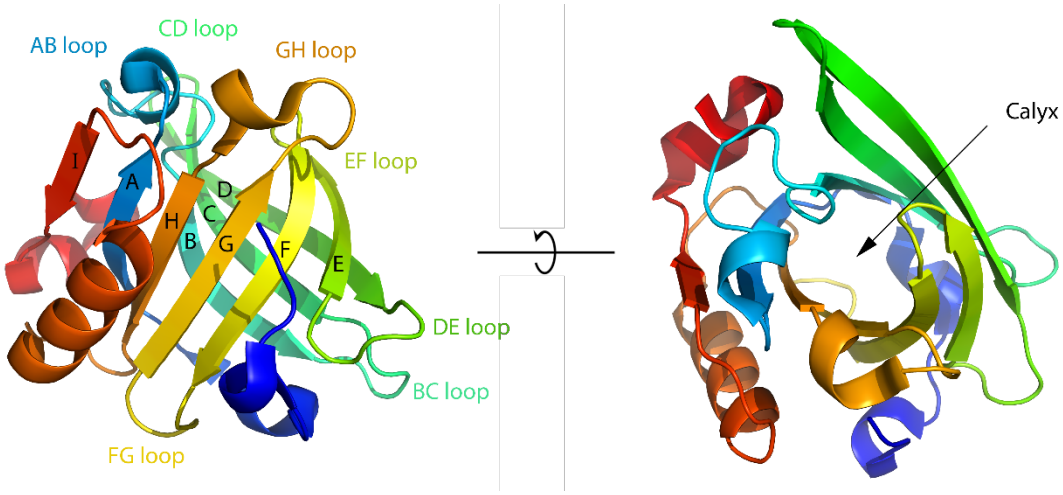


Figure 4.8: Crystal structure of one monomer of caprine βlg (PDB ID: 4TLJ (Crowther *et al.* 2014)) as seen from the side and from above. β-strands are labelled A to I. The polypeptide chain is shown in rainbow colours, beginning in blue at the N-terminus and ending in red at the C-terminus. (Crowther *et al.*, 2016) © under CC BY 3.0 license.

The loops at the closed end of the calyx (BC, DE and FG) are quite short, whereas those at the open end (AB, CD, EF and GH) are longer and more flexible (Figure 4.8) (Jameson *et al.*, 2002). The long CD and GH loops are quite flexible in both caprine and bovine βlg structures, however the electron density is clearly defined for caprine βlg, in contrast to bovine βlg structures. Furthermore, whereas the C-terminal region is ill-defined in bovine structures, it is well defined here along with improvement in definition of the AB loops at the dimer interface and the pH-gated EF loop.

The EF loop, which is present at the mouth of the calyx and undergoes a pH induced conformational transition in bovine βlg (Qin *et al.*, 1998b), is in the closed conformation in this structure determined at pH ~6.8 (Figure 4.9). This is consistent with findings for bovine βlg, which show the EF loop in a closed conformation at pH 6.5, burying Glu89 as glutamic acid,

and in an open conformation at pH values above 7, exposing Glu89 as a glutamate (Qin *et al*., 1998a). A recent structure of caprine βlg at 2.3 Å and pH 7.25 (published during the course of this study) (Loch *et al*., 2015) shows an asymmetric dimer with the EF loop of one subunit in the closed position and the other in the open position. This suggests that caprine βlg is also capable of undergoing this conformational transition.

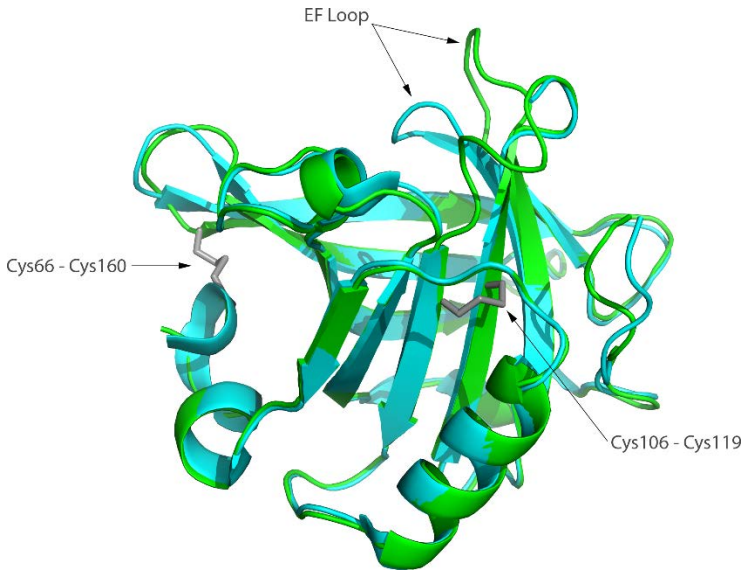


Figure 4.9: The crystal structure of caprine βlg (blue, PDB ID: 4TLJ) overlaid with the structure of bovine βlg (green, PDB ID: 1BSY (Qin *et al.*, 1998a)) (rmsd = 0.397 Å). The EF loop is open in the bovine βlg structure crystallised at pH 7.1, while it is closed in the caprine βlg structure crystallised at pH 6.8. βLg from both species contain two disulfide bonds per monomer (shown as grey sticks) and a single free thiol.

Caprine βlg shares the same arrangement of disulfide bonds as seen in bovine, ovine and reindeer structures of βlg. Each monomer of βlg contains five cysteine residues. One exists as a free thiol on strand G (Cys121) and is buried beneath the α-helix that lies alongside the β-barrel, whereas the other four form two disulfide bridges. The first (Cys66-Cys160) links the C-terminus to the CD loop, while the second (Cys106-Cys119) links β-strands G and H (Figure 4.9). As discussed in Chapter 3, the presence and correct arrangement of these disulfide bonds are crucial for the correct folding of bovine βlg and are believed to contribute to the acid stability of this protein (Ponniah *et al*., 2010). Given their presence within caprine βlg, it comes

as little surprise that the caprine orthologue also exhibits pronounced acid stability and can thus be purified using the same protocol as bovine βlg (see section 3.2.2).

4.2.3 Solution structure of bovine and caprine βlg

4.2.3.1 Small angle X-ray scattering analysis

Small-angle X-ray scattering (SAXS) data indicate that the crystal structure of caprine βlg is an accurate representation of the protein structure in solution. SAXS is a lower resolution technique than X-ray crystallography, yet its advantage lies in the fact that the data is collected from molecules in solution. The scattering profiles obtained provide information about the size, shape and orientation of molecules within a sample. This makes it a useful complement to X-ray crystallography as it can indicate whether a model built from diffraction of a crystallised protein is a true representation of that protein's structure in solution.

Small angle X-ray scattering data were collected for samples of the A and B variants of bovine βlg and caprine βlg (Figure 4.10). At small values of q (the scattering vector), the linear portion of the Guinier plot (ln(I) vs q^2) can be used to extrapolate back to I(0) (scatter intensity at the origin) and to estimate R_g (the radius of gyration). The linearity of the Guinier plots suggests that the samples were free from aggregation or inter-particle effects.

An indirect Fourier transform, performed using GNOM, yields the function $P(r)$, which gives the relative probabilities of distances between scattering centres and the maximum dimension of the particle, D_{max}. For all three protein samples, the maximum dimension of the particle was in agreement, at ~72 Å (Figure 4.10 C). The rather asymmetrical shape of the $P(r)$ plots and the tail at high q values indicate that these proteins are dumbbell-shaped (i.e. reasonably globular with elongation in one direction). I(0) and R_g can also be derived from the $P(_r)$ distribution and comparison of these parameters with those estimated from the Guinier analysis provides a useful cross-check. In this case, the parameters agree well. The structural parameters are summarised in Table 4.4.

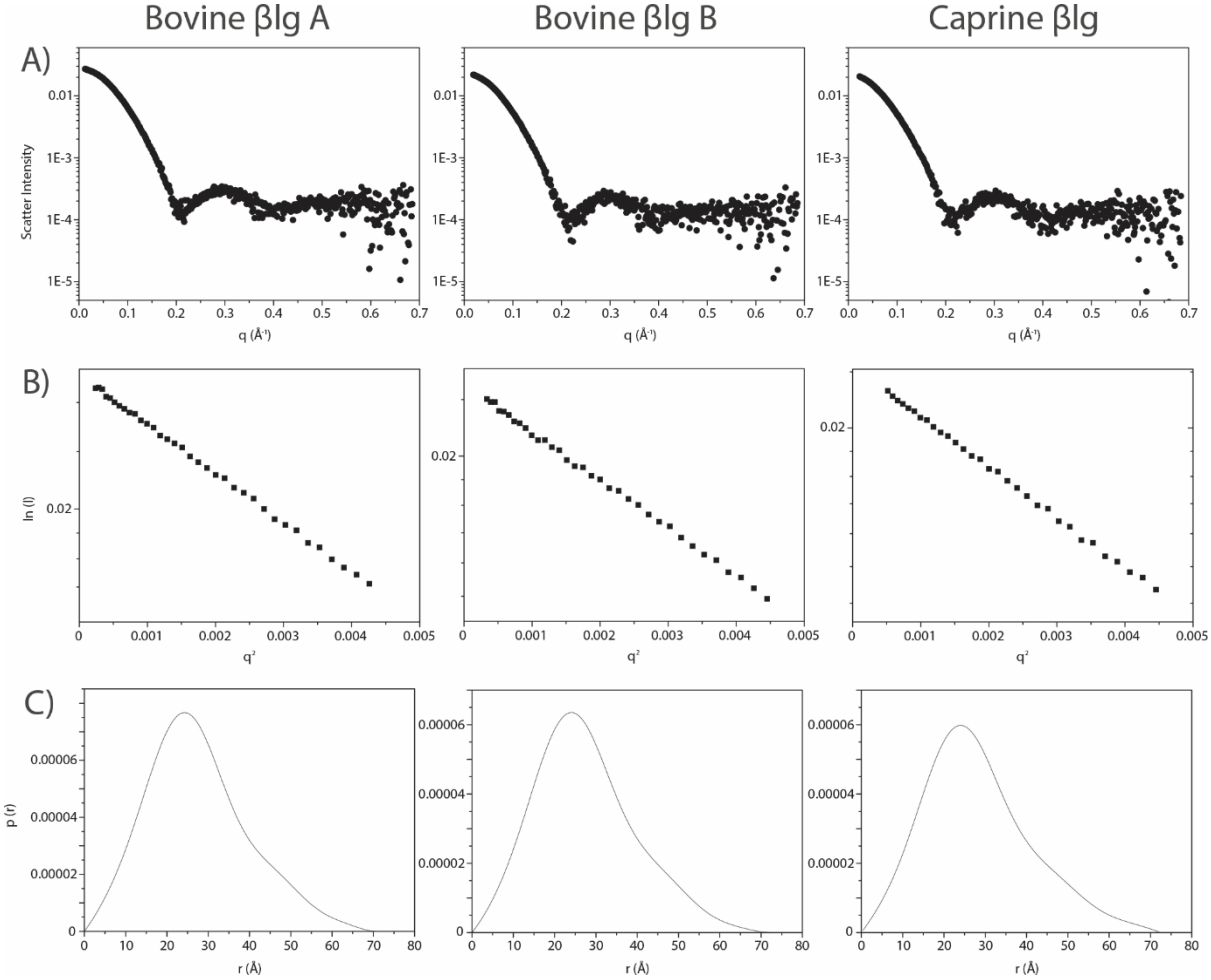


Figure 4.10: A) Small-angle X-ray scattering profiles. B) Guinier plots. C) *P*(*r*) distributions.

Table 4.4: Small-angle X-ray scattering parameters

	Guinier		Gnom			
	R_g (Å)	I(0)	R_g (Å)	I(0)	Porod Volume	D_{max} (Å)
Bovine βlg A	21.16	0.027	21.37	0.03	40731.2	70.91
Bovine βlg B	21.03	0.023	21.33	0.02	42104.1	72.52
Caprine βlg	21.75	0.022	21.93	0.02	44638.0	72.59

4.2.3.2 Scattering pattern calculation from atomic models

Due to the nature of crystallisation, the crystal structure of a protein may not always reflect the structure of the protein when in solution. In order to assess whether a crystal structure is an accurate representation of that protein in solution, the scattering that would result from a protein with that exact structure can be calculated. This theoretical scattering profile can then be compared to the experimental scattering profile.

The theoretical scattering profiles that would result from crystal structures of bovine and caprine βlg were calculated for both an individual subunit and for a dimeric arrangement of subunits. The theoretical scattering profiles were compared to the experimental scattering collected for bovine βlg A and B and caprine βlg. In each case, the experimental scattering was a much closer fit to the theoretical scattering of a dimer of βlg than to a monomer of βlg (Figure 4.11). This indicates that crystal structures of βlg are an accurate reflection of the protein structure in solution. The fits to the theoretical scattering profiles of βlg dimers are still not perfect, which may be due to flexibility in the arrangement of subunits within the dimer. This is discussed in more detail in section 4.2.4 and when assessing the dynamic nature of βlg in Chapter 5 (section 5.2.2).

The closer agreement between the theoretical scattering profiles of βlg dimers with the experimental scatter indicates that each βlg variant is predominantly a dimer at the concentrations examined. Given that the proteins were initially loaded at ~10 mg/mL before being subjected to size-exclusion, which results in approximately a 10-fold dilution, the resultant protein concentration is likely to be ~1 mg/mL, or ~55 μM. These findings reinforce the observations made from analytical ultracentrifugation data that each of the βlg variants are predominantly dimeric at near-neutral pH and at concentrations above 50 μM.

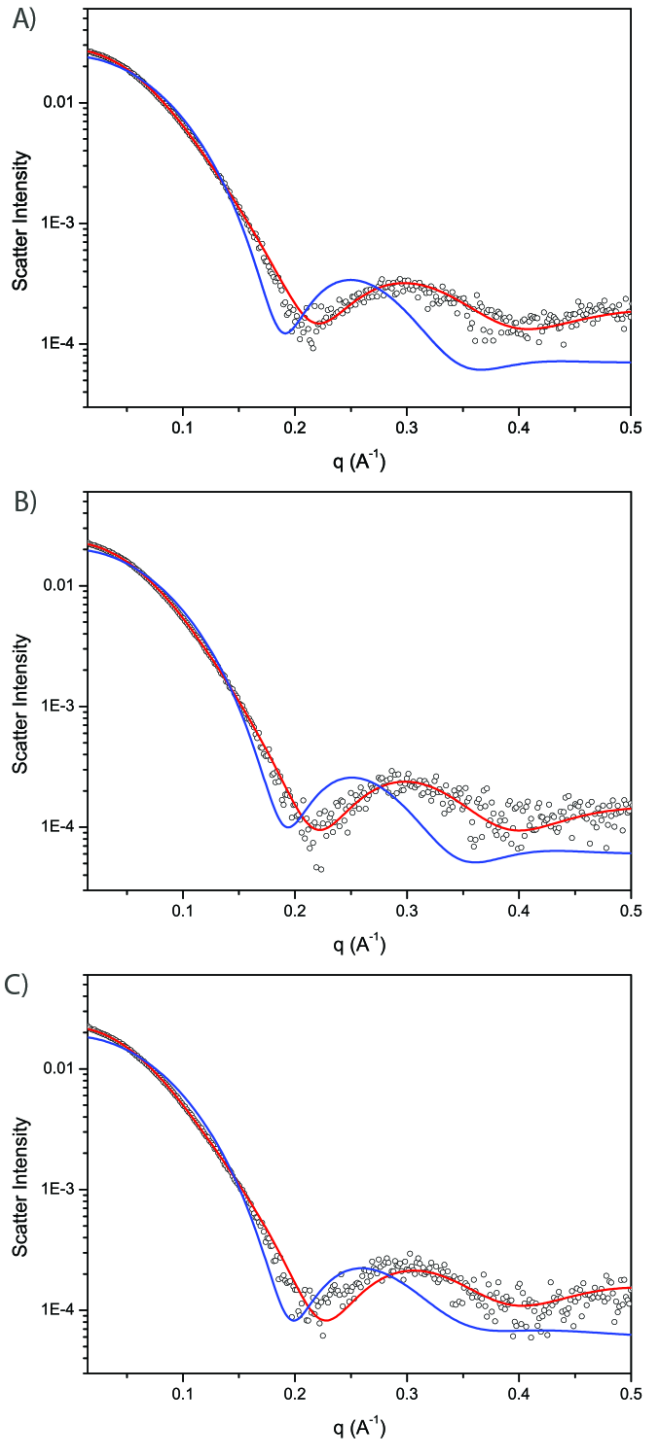


Figure 4.11: Theoretical scattering profiles for the dimer (red line) and the monomer (blue line) from βlg crystal structures, overlaid with experimental SAXS data (o). Only one representative PDB structure for each bovine variant is shown as similar results were seen for other structures of bovine βlg. A) Bovine βlg A scattering data with curves generated from PDB ID: 1BSY (Qin *et al.*, 1998a), χ^2 = 0.185 (dimer) and 9.037 (monomer). B) Bovine βlg B scattering data with curves generated from PDB ID: 1BSQ (Qin *et al.*, 1999), χ^2 = 0.145 (dimer) and 6.826 (monomer). C) Caprine βlg scattering data with curves generated from PDB ID: 4TLJ, χ^2 = 0.195 (dimer) and 9.752 (monomer).

4.2.4 Comparison of βlg orthologue structures

Crystal structures of bovine, ovine (sheep), reindeer and here caprine βlg indicate that these orthologues all share a high degree of structural similarity (Figure 4.12). However, there are significant differences between the structures of these orthologues and that of porcine (pig) βlg. This is not unreasonable given the lower level of sequence identity (63 – 65%) between porcine βlg and bovine, ovine, caprine and reindeer βlg, where the latter four share sequence identity in pair-wise comparisons of 93 – 99% (Table 4.5). While the monomers share a very similar secondary structure arrangement, porcine βlg features a completely different quaternary association.

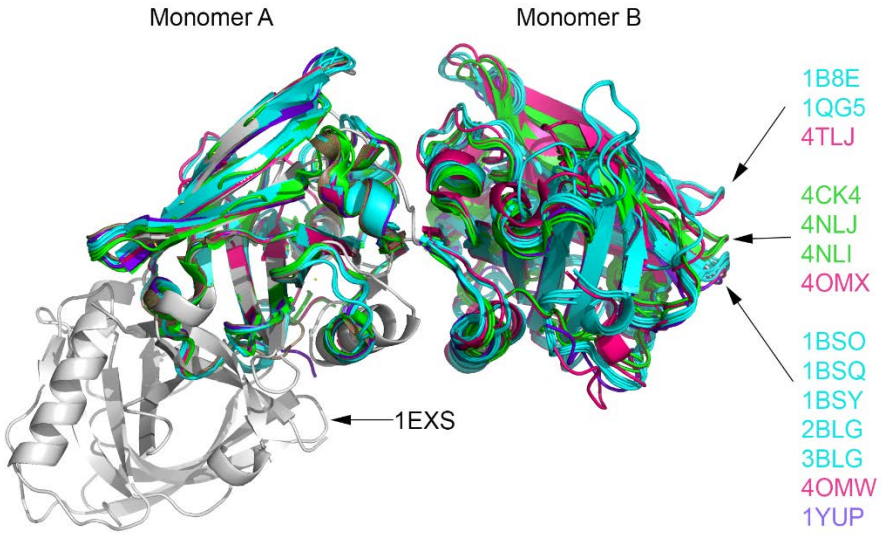


Figure 4.12: Superimposed crystal structures of caprine βlg (PDB: 4TLJ (Crowther *et al.* 2014), 4OMX, 4OMW (Loch *et al.*, 2015) = pink), bovine βlg A (PDB: 1QG5 (Oliveira *et al.*, 2001), 2BLG, 3BLG, 1BSY (Qin *et al.*, 1998a), 1BSO (Qin *et al.*, 1998b) = cyan), bovine βlg B (PDB: 1B8E (Oliveira *et al.*, 2001), 1BSQ (Qin *et al.*, 1999) = cyan), ovine βlg (PDB: 4CK4 (Kontopidis *et al.*, 2014), 4NLI, 4NLJ (Loch *et al.*, 2014) = green), reindeer βlg (1YUP (Oksanen *et al.*, 2006) = purple) and porcine βlg (1EXS (Hoedemaeker *et al.*, 2002) = grey). Monomer A of each structure is aligned to highlight the variation in orientation of the second monomer. A table of all PDB X-ray crystallography protein structures referenced throughout this thesis is included in the appendix.

While there is a high level of structural similarity between the caprine, bovine, ovine and reindeer orthologues of βlg, there are differences in the way the protein subunits assemble. The rmsd for the superposition of Cα atoms of these structures as dimers onto the structure of caprine βlg solved here (4TLJ) ranges from 0.512 to 1.534 Å (Table 4.5). This range is due to the variable orientation of the second monomer in the dimer with respect to the first (Figure 4.12). When a single monomer is used for the alignment the range in rmsd reduces to 0.347 to 0.551 Å (Table 4.5). The small rmsd values that result from aligning monomers indicate that the tertiary fold of βlg is similar between closely related orthologues, but that there is flexibility in the orientation of the monomers relative to each other.

The varied subunit arrangement seen in crystal structures of orthologues of βlg is not correlated with the position of the EF loop as the structures selected for comparison include pH values either side of the pH at which this conformational change occurs (~pH 7) (3BLG: pH 6.1, 4TLJ: pH 6.8, 1BSY and 1BSQ: 7.1, 1BSO: 7.3, 1B8E and 1QG5: 7.9, and 2BLG: 8.2). The selection also includes both bovine variants A and B (1QG5: bovine variant A and 1B8E: bovine variant B), and one ligand-bound bovine βlg structure (1BSO). Alternatively, the varied arrangement of monomers in the dimer may be an artefact of crystal packing as structures belonging to the same space groups are the most closely related, for instance the bovine βlg structures 1QG5 and 1B8E (space group $C222_1$) have a similar monomer arrangement to each other, whereas the bovine βlg structures 1BSQ, 2BLG, 1BSY, 1BSO and 3BLG (space group $P3_221$) share a different arrangement of monomers (Figure 4.13).

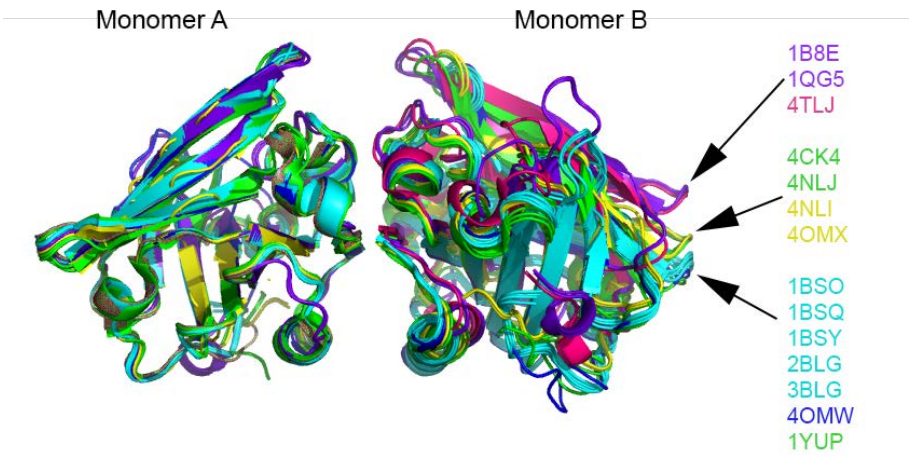


Figure 4.13: The same crystal structures as in Figure 4.12 except the models are coloured according to space group: $P2_12_12$ = blue, $P3_221$ = cyan, P1 = green, $P3_121$ = yellow, $P12_11$ = pink, $C222_1$ = purple.

Table 4.5. The percentage sequence identity of βlg orthologue protein sequences, the rmsd when aligning the Cα atoms of monomers and the rmsd when aligning the Cα atoms of dimers. Alignments all performed against the structure of caprine βlg (4TLJ) in PyMol, aligning residues 1 – 162.

Species	Sequence identity (%)	Structure	Monomer rmsd (Å)	Dimer rmsd (Å)
Caprine	100	4OMW	0.368	1.534
		4OMX	0.347	0.805
Bovine A	95.06	1QG5	0.456	0.541
		2BLG	0.444	1.332
		3BLG	0.376	1.524
		1BSY	0.397	1.4
		1BSO	0.409	1.513
Bovine B	96.3	1B8E	0.44	0.512
		1BSQ	0.386	1.318
Ovine	99.38	4CK4	0.551	1.076
		4NLJ	0.52	0.745
		4NLI	0.359	0.776
Reindeer	93.21	1YUP	0.372	1.373
Porcine	62.5	1EXS	1.398	19.689

4.2.4.1 Amino acid substitutions between bovine and caprine βlg

The amino acid substitutions between caprine βlg and the bovine variants of βlg do not appear to have any significant effects on the protein structure (Figure 4.14). The substitutions L1I and I162V occur at the N- and C-termini respectively and are reasonably conservative changes that would be expected to have a limited effect on protein structure. While the position of the N-terminus is different for caprine βlg than for bovine βlg A and B in the structures chosen for comparison, this is most likely due to the flexibility of this region. The representative structures of bovine βlg A and B (PDB ID's 1QG5 and 1B8E, respectively) were chosen for illustration due to them sharing the most similar arrangement of monomers with caprine βlg, as seen in Figure 4.12.

The substitution E158G is a more significant change, as it results in the loss of a sidechain carboxylic acid. However, this substitution occurs near the C-terminus, which is tethered in place by the disulfide bond Cys66-160, thus this substitution is likely to have little effect on the overall structure at this site. In the structures of bovine A and B the C-termini from residues 154 and 152, respectively, are ill-defined and have not been modelled. In contrast the C-terminus is well defined for caprine βlg.

The other three substitutions have the potential to exert a greater effect on the structure of caprine βlg. D53N involves the substitution of a negatively charged (at neutral pH) aspartate residue for a polar asparagine residue. This appears to have only a slight effect on the position of the nearby BC loop, however its contribution to the difference in net charge on the molecule may have significance as it has the potential to affect the way this protein interacts with other components in milk.

D130K is another substitution involving charged side-chains. In this case the negatively charged aspartate residue (at neutral pH) is swapped for a positively charged lysine. This residue is located on the α-helix near the dimer interface. While this substitution does not appear to alter the protein structure, it has the potential to affect the oligomerisation behaviour, particularly at low pH when the positively charged monomers begin to repel each other, as seen in section 4.2.1. This is discussed in more detail in section 4.2.4.2.

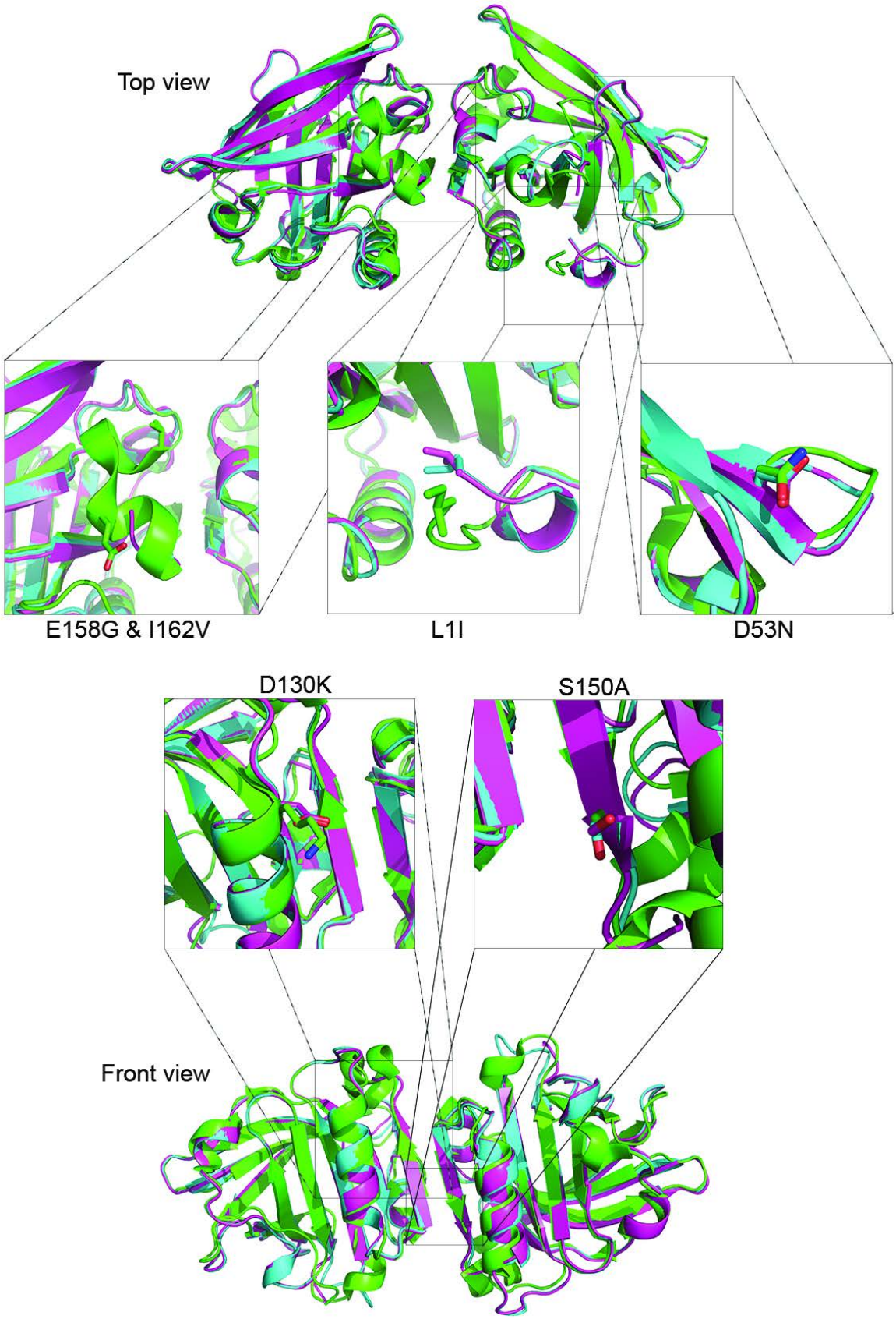


Figure 4.14: Amino acid substitutions between caprine βlg (PDB ID: 4TLJ, green) and bovine A and B βlg (PDB ID's: 1QG5 and 1B8E (Oliveira *et al.*, 2001), pink and blue).

The final substitution, S150A, results in the loss of a hydroxyl sidechain. While this is near the dimer interface and could therefore perceivably affect dimer formation, this is unlikely as in crystal structures of bovine βlg the side chain of this serine is not close enough to any other residues to form any inter-molecular hydrogen bonds. While none of these six substitutions appear to significantly alter caprine βlg's structure, it is still possible that they have altered the protein's dynamics. This is investigated further in Chapter 5.

4.2.4.2 Differences in quaternary association

The crystal structure of caprine βlg suggests that it shares the same dimer interface as seen in structures of bovine βlg, including the hydrogen bonds between I-strands and the salt bridges between AB loops (Figure 4.2). The existence of these elements in the caprine βlg structure suggests that this protein will exhibit similar oligomerisation behaviour as bovine βlg. However, caprine βlg displays a dependence on pH of the quaternary association that differs from bovine βlg.

Can the difference in oligomerisation behaviour at low pH be rationalised in terms of the sequence differences between the bovine and caprine orthologues? Both bovine βlg A and B have three lysines on the α-helix that runs along the outside of the β-barrel (Lys-135, Lys-138 and Lys-141) (Figure 4.15). There are also several negatively charged aspartic acid and glutamic acid residues along this helix that partly neutralise the charge of these positive lysines. As the pKa values of aspartic and glutamic acid residues are usually around 4, these residues lose their negative charge when the pH drops below this value. This leaves an area of positive charge near the dimer interface at low pH, which causes dissociation of the dimer, unless there are enough negative ions in solution to shield these charges (Mercadante *et al*., 2011). In caprine βlg there is an additional positively charged lysine on this α-helix, in the place of a negatively charged aspartic acid in bovine βlg (D130K) (Figure 4.15). Therefore, at low pH there is a larger area of positive charge in the vicinity of these α-helices. This may mean that a greater ionic strength is required to stabilise dimer formation of caprine βlg at low pH.

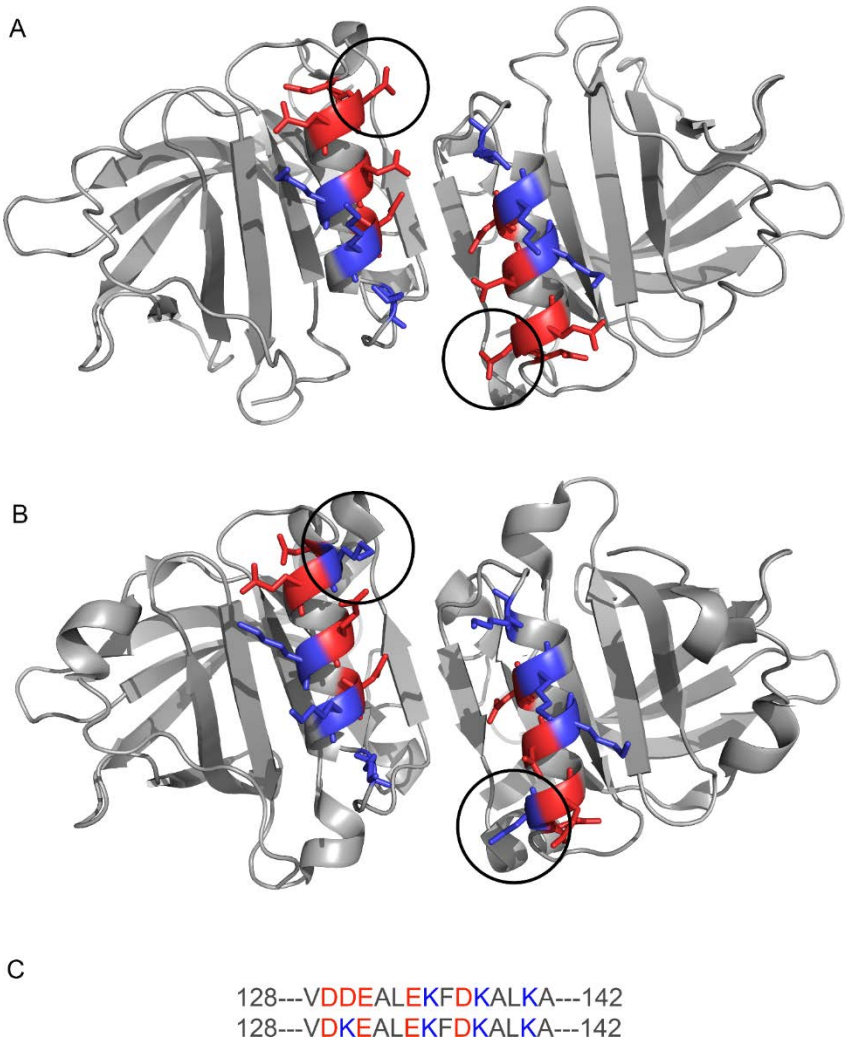


Figure 4.15: The D130K substitution in caprine βlg swaps a negatively charged residue for a positively charged residue on the α-helix near the dimer interface. A) Bovine βlg. B) Caprine βlg. C) Protein sequence between residues 128 and 142 of bovine βlg versus caprine βlg. Lysines are shown in blue, and aspartic acid and glutamic acid residues are shown in red.

4.2.4.3 The epitope binding site of βlg

The structure of bovine βlg in complex with an IgE fragment (Niemi *et al*., 2007) reveals the location of the epitope binding site of bovine βlg (i.e. the surface region where IgE interacts with bovine βlg). Despite the fact that there are four amino acid substitutions between bovine and caprine βlg located near the epitope region (D53N, D130K, E158G, I162V), there is very little difference in the structure of caprine βlg at the corresponding site of the epitope of bovine βlg (Figure 4.16). Therefore, it is likely that IgE antibodies raised against bovine βlg will

also be capable of interacting with caprine βlg. However, it is possible that the amino acid substitutions between caprine and bovine βlg may alter the protein dynamics, which may then affect the affinity of IgE antibodies. The dynamic behaviour of these orthologues is investigated in Chapter 5. Further, if caprine βlg is digested differently than bovine βlg it may be presented to the immune system in an altered form which may no longer be cross-reactive. The comparative stability of bovine and caprine βlg is also investigated in Chapter 5.

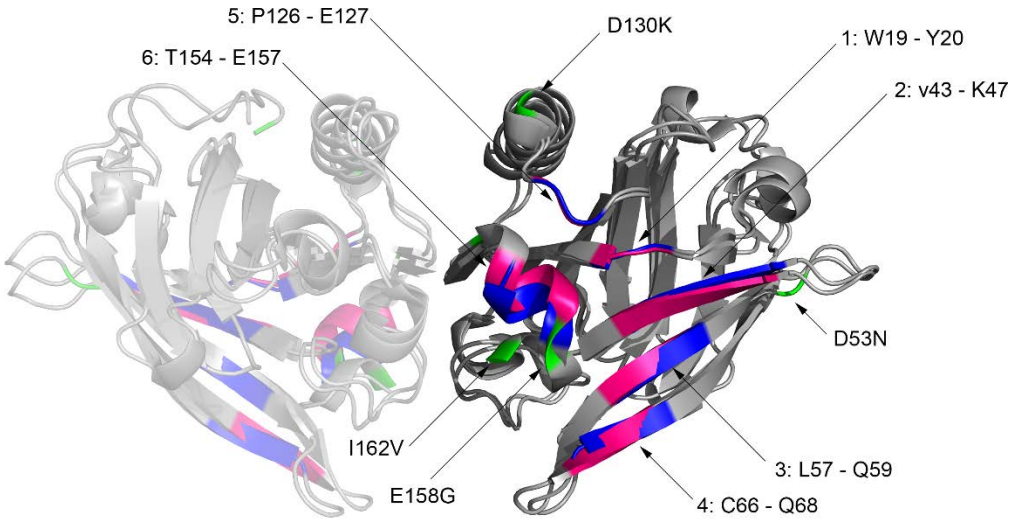


Figure 4.16: Aligned structures of bovine and caprine βlg (PDB ID: 2R56 (Niemi *et al.,* 2007) and 4TLJ (Crowther *et al.*, 2014)). The residues of bovine βlg that make contact with IgE are shown in blue. The corresponding residues in caprine βlg are shown in pink. The different segments of the epitope are labelled 1 – 6. The amino acid substitutions between bovine and caprine βlg are shown in green.

4.3 Summary

In order to assess the quaternary state association of caprine βlg a sedimentation velocity analysis was carried out. This analysis followed the protocol of that of Mercadante *et al*. (2011) so that the results could be compared to those of bovine βlg. The results showed that at pH 2.5 caprine βlg is predominantly monomeric between concentrations of 5 and 50 μM, whereas bovine βlg is in a monomer-dimer self-association under the same conditions. The substitution of an aspartic acid residue for a lysine at position 130, located on an α-helix in close vicinity to

the dimer interface, potentially influences the oligomerisation behaviour of caprine βlg at low pH. The added positive charge may mean that a greater ionic strength is required to shield charge repulsion between positively charged subunits at low pH in order to stabilise dimer formation.

Elucidation of an ultra-high resolution crystal structure has allowed the structure of caprine βlg to be modelled for the first time. This confirms that the caprine orthologue of βlg shares a similar architecture with that of the closely related bovine, ovine and reindeer orthologues. Further, the ultra-high resolution of this new structure has made it possible to clearly define features that in lower resolution bovine structures are obscured by disorder. While sequence substitutions between bovine and caprine βlg do not appear to have a significant influence on the local structure, it is possible that these changes have altered the protein dynamics.

Small-angle X-ray scattering and analytical ultracentrifugation analyses confirm that the solution structure of caprine and bovine βlg is well represented in crystal structures of these proteins. Between pH 6.5 and 7.5 bovine and caprine βlg self-associate into dimers with dissociation constants within the range of the concentrations studied (5 - 50 μM). Both of these proteins are, therefore, likely to be dimeric at the concentration and pH likely to be encountered within milk (~ 3 mg/mL or ~150 μM and pH 6.5).

The structure of bovine βlg in complex with an IgE antibody fragment supports claims that βlg contributes to cow milk allergies. Goat milk has a reduced allergenic burden compared to cow milk (Bellioni-Businco *et al*., 1999; Lara-Villoslada *et al*., 2004). This could be the result of a lower affinity of antibodies for goat milk allergens, perhaps due to structural differences. The elucidation of the structure of caprine βlg indicates that the fold of caprine βlg is very similar to that of bovine βlg proteins.

Alternatively it may be that goat milk allergens are presented to the immune system in a different form such that they do not present the same target to IgE antibodies. This could be due to differences in quaternary structure, alterations to protein structure caused by processing or due to differences in digestibility. The difference in the oligomerisation behaviour at low pH between caprine and bovine βlg has the potential to affect how these proteins are presented to the immune system during digestion. It is necessary to assess the

stability and dynamics of these proteins in order to gain a greater understanding of how these proteins will behave during processing and digestion. These features are addressed in Chapter 5.

4.4 References

Almaas, H., Holm, H., Langsrud, T., Flengsrud, R., Vegarud, G.E., 2006. *In vitro* studies of the digestion of caprine whey proteins by human gastric and duodenal juice and the effects on selected microorganisms. The British journal of Nutrition 96, 562–9.

Bellioni-Businco, B., Paganelli, R., Lucenti, P., 1999. Allergenicity of goat's milk in children with cow's milk allergy. Journal of Allergy and Clinical Immunology 1191–1194.

Bello, M., del Portillo-Téllez, M., García-Hernández, E., 2011. Energetics of Ligand Recognition and Self-Association of Bovine β-Lactoglobulin: Differences between Variants A and B. Biochemistry-us 50, 151–161.

Brownlow, S., Cabral, J., Cooper, R., Flower, D., Yewdall, S., Polikarpov, I., North, A., Sawyer, L., 1997. Bovine β-lactoglobulin at 1.8 Å resolution — still an enigmatic lipocalin. Structure 5, 481–495.

Crowther, J., Lassé, M., Suzuki, H., Kessans, S., Loo, T., Norris, G., Hodgkinson, A., Jameson, G., Dobson, R., 2014. Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin. FEBS Letters 588, 3816–3822.

Crowther, J., Jameson, G., Hodgkinson, A., Dobson, R., 2016. Structure, Oligomerisation and Interactions of β-Lactoglobulin. InTech.

Hoedemaeker, F.J., Visschers, R.W., Alting, A.C., de, K.G., Kuil, M.E., Abrahams, J.P., 2002. A novel pH-dependent dimerization motif in β-lactoglobulin from pig (Sus scrofa). Acta Crystallogr Sect D Biological Crystallogr 58, 480–486.

Jameson, G.B., Adams, J.J., Creamer, L.K., 2002. Flexibility, functionality and hydrophobicity of bovine β-lactoglobulin. International Dairy Journal 12, 319–329.

Kapila, R., Kavadi, P., Kapila, S., 2013. Comparative evaluation of allergic sensitization to milk proteins of cow, buffalo and goat. Small Ruminant Research 112, 191–198.

Kobayashi, T, Ikeguchi, M, Sugai, S, 2002. Construction and characterization of β-lactoglobulin chimeras. Proteins 49, 297–301.

Kontopidis, G, Holt, C, Sawyer, L, 2004. Invited review: β-lactoglobulin: binding properties, structure, and function. Journal of Dairy Science 87, 785–796.

Kontopidis, G., Gilliver, A., Sawyer, L., 2014. Ovine β-lactoglobulin at atomic resolution. Acta Crystallographica Section F: Structural Biology Communications 70, 1498–1503.

Lara-Villoslada, F., Olivares, M., Jiménez, J., Boza, J., Xaus, J., 2004. Goat milk is less immunogenic than cow milk in a murine model of atopy. Journal of Pediatric Gastroenterology and Nutrition 39, 354–360.

Laue, T.M., Stafford, W.F., 1999. Modern applications of analytical ultracentrifugation. Annual Review of Biophysics and Biomolecular Structure 28, 75–100.

Le Maux, S., Bouhallab, S., Giblin, L., Brodkorb, A., Croguennec, T., 2014. Bovine β-lactoglobulin/fatty acid complexes: binding, structural, and biological properties. Dairy Science & Technology 94, 409–426.

Lebowitz, J., Lewis, M., Schuck, P., 2002. Modern analytical ultracentrifugation in protein science: A tutorial review. Protein Science 11, 2067–2079.

Loch, J., Molenda, M., Kopeć, M., Świątek, S., Lewiński, K., 2014. Structure of two crystal forms of sheep β-lactoglobulin with EF-loop in closed conformation. Biopolymers 101, 886–894.

Loch, J., Bonarek, P., Polit, A., Świątek, S., Czub, M., Ludwikowska, M., Lewiński, K., 2015. Conformational variability of goat β-lactoglobulin: Crystallographic and thermodynamic studies. International Journal of Biological Macromolecules 72, 1283–1291.

McKenzie, HA, Sawyer, WH, 1972. On the dissociation of bovine beta-lactoglobulins A, B, and C near pH 7. Australian Journal of Biological Sciences 25, 949–961.

Mercadante, D., Melton, L., Norris, G., Loo, T., Williams, M., Dobson, R., Jameson, G., 2011. Bovine β-Lactoglobulin Is Dimeric Under Imitative Physiological Conditions: Dissociation Equilibrium and Rate Constants over the pH Range of 2.5–7.5. Biophysical Journal 103, 303–312.

Niemi, M., Jylhä, S., Laukkanen, M.-L., Söderlund, H., Mäkinen-Kiljunen, S., Kallio, J., Hakulinen, N., Haahtela, T., Takkinen, K., Rouvinen, J., 2007. Molecular Interactions between a Recombinant IgE Antibody and the β-Lactoglobulin Allergen. Structure 15, 1413–1421.

Ohtomo, H., Fujiwara, K., Ikeguchi, M., 2012. Important role of methionine 145 in dimerization of bovine β-lactoglobulin. Journal of biochemistry 151, 329–34.

Ohtomo, H., Konuma, T., Utsunoiya, H., Tsuge, H., Ikeguchi, M., 2011. Structure and stability of Gyuba, a β-lactoglobulin chimera. Protein Science 20, 1867–1875.

Oksanen, E., Jaakola, V.P., Tolonen, T., Valkonen, K., Åkerström, B., Kalkkinen, N., Virtanen, V., Goldman, A., 2006. Reindeer β-lactoglobulin crystal structure with pseudo-body-centred noncrystallographic symmetry. Acta Crystallographica Section D: Biological Crystallography 62, 1369–1374.

Oliveira, K.M., Valente-Mesquita, V.L., Botelho, M.M., Sawyer, L., Ferreira, S.T., Polikarpov, I., 2001. Crystal structures of bovine β-lactoglobulin in the orthorhombic space group C2221. European Journal of Biochemistry 268, 477–484.

Ponniah, K., Loo, T., Edwards, P., Pascal, S., Jameson, G., Norris, G., 2010. The production of soluble and correctly folded recombinant bovine β-lactoglobulin variants A and B in *Escherichia coli* for NMR studies. Protein Expression and Purification 70, 283–289.

Qin, B., Bewley, M., Creamer, L., Baker, H., Baker, E., Jameson, G., 1998a. Structural Basis of the Tanford Transition of Bovine β-Lactoglobulin. Biochemistry 37, 14014–14023.

Qin, B., Creamer, L., Baker, E., Jameson, G., 1998b. 12-Bromododecanoic acid binds inside the calyx of bovine β-lactoglobulin. FEBS Letters 438, 272–278.

Qin, B., Jameson, G., Bewley, M., Baker, E., Creamer, L., 1999. Functional implications of structural differences between variants A and B of bovine β-lactoglobulin. Protein Science 8, 75–83.

Rhodes, G., 2006. Crystallography Made Crystal Clear, 3rd ed. Elsevier.

Sakurai, K., Goto, Y., 2002. Manipulating Monomer-Dimer Equilibrium of Bovine β-Lactoglobulin by Amino Acid Substitution. Journal of Biological Chemistry 277, 25735–25740.

Sakurai, K., Oobatake, M., Goto, Y., 2001. Salt-dependent monomer–dimer equilibrium of bovine β-lactoglobulin at pH 3. Protein Sci 10, 2325–2335.

Chapter Five

5. Stability and Dynamics of Caprine β-lactoglobulin

5.1 Introduction

5.1.1 Caprine βlg is digested more efficiently than bovine βlg

The nutritional value of proteins is dependent on the bioavailability of their amino acid components. Bovine βlg is resistant to peptic and chymotryptic digestion and thus remains largely intact after passing through the digestive tract (Reddy *et al*., 1988). The resistance to proteolytic digestion may be linked to the allergenicity of bovine βlg, given that much of this allergen passes through the digestive tract in its native form.

Compared to bovine βlg, caprine βlg is much more efficiently broken down during the *in vitro* digestion of cow and goat milk (Almaas *et al*., 2006). After treatment with human gastric juice and duodenal juice in a simulated digestion experiment, 83% of bovine βlg remained undigested in cow milk, while in goat milk only 23% of caprine βlg remained undigested (Almaas *et al*., 2006). The smaller size of the fat globules and their natural dispersion in goat milk allows for greater access by lipases (Almaas *et al*., 2006b) , while the reduced α_{s1}-casein protein content leads to a softer curd formation in the gut, which allows better access by proteases (Ceballos *et al*., 2009). These factors lead to enhanced fat and protein digestion in goat milk. However, it remains to be seen whether the differential digestibility of bovine and caprine βlg proteins are a reflection of the different composition of these milks, or is in fact due to the stability of these proteins themselves.

5.1.2 Cow and goat milk respond differently to heat treatment

As discussed in section 1.5, it has been found that goat milk is less heat stable than cow milk at a given pH (Anema and Stanley, 1998; Bouhallab *et al*., 2002; Montilla and Calvo, 1997; Zadow *et al*., 1983) . The heat stability of cow and goat milk is inherently linked to the association of denatured whey proteins, such as βlg, with κ-casein, and whether these complexes then dissociate or stay attached to casein micelles. This interaction behaviour is affected by the pH of the milk, the structure of the casein micelles, the concentration of ions such as calcium and phosphate, and by changes in protein interactions due to variations in the primary structures of these proteins. What is not known is whether the thermal stability of the individual proteins is at all responsible for the different heat treatment responses seen between cow and goat milk.

5.1.3 The different behaviours of bovine and caprine βlg may be due to the stability and dynamics of these proteins

Conformational stability is the result of various forces (van der Waals interactions, hydrogen bonds, electrostatic interactions/salt-bridges and disulfide bonds) that work to keep a protein in its folded form. When elements of the protein's environment are altered, for instance the temperature or the pH, or if chemical denaturants are added, the protein may unfold and become denatured. This is due to the disruption of hydrogen bonds and other interactive forces responsible for holding together the secondary, tertiary and quaternary structure of proteins. Due to the differences in the primary structure of caprine and bovine βlg proteins (8 amino acid differences between caprine βlg and bovine βlg A and 6 between caprine βlg and bovine βlg B) it is conceivable that these proteins may have different conformational stabilities, as the interactions holding these proteins together may be altered. A difference in conformational stability may lead to a difference in digestibility as the backbone of the protein may be more susceptible to attack by digestive proteolytic enzymes. It may also alter the responses of these proteins towards heat treatment.

While the changes in primary structure between these βlg orthologues have little effect on the secondary and tertiary structure (section 4.2.4), there is evidence that these changes affect the quaternary association of these proteins (section 4.2.1). It is also possible that these changes may influence the dynamic nature of these proteins. Regions of proteins that are accessible to proteases are susceptible to proteolysis during digestion. Increased flexibility of a protein may increase the accessibility for proteases, thus increasing the protein's digestibility.

5.1.4 Overview

The aim of this chapter is to characterise and compare the stability of these proteins toward urea and thermal denaturation. This information is necessary to understand the different behaviours exhibited by cow and goat milk, and to understand the response of these milks to processing conditions and digestion.

All three proteins (bovine βlg A and B and caprine βlg) exhibit similar resistance towards unfolding by urea, indicating a similar level of conformational stability. Their responses to thermal denaturation, however, show a markedly different dependence on pH. While the melting temperature of the bovine βlg variants increases as the pH decreases (from pH 7.4 to 6.5), caprine βlg appears to be very sensitive to thermal denaturation at pH values below 6.9 and considerable precipitation is observed. The dependence on pH of the heat-stability of this protein parallels with that of goat milk, suggesting that this behaviour in milk may have more to do with βlg itself than previously proposed.

As outlined in chapter 4, the structures of bovine and caprine βlg are very similar. The changes in the primary sequence of these proteins may instead alter the dynamics of these proteins, which in turn may be responsible for their different properties. Thus, the dynamics of these proteins were simulated computationally. Molecular dynamics simulations show that the same regions of each βlg protein have the largest fluctuations, namely the flexible loops found at each end of the β-barrel. Interestingly, significant subunit rearrangement and even brief dimer dissociation events were observed. There appears to be considerable flexibility at the dimer interface and numerous interactions are observed that are not found in crystal structures of βlg.

5.2 Results and discussion

5.2.1 Assessing the conformational stability of βlg proteins

5.2.1.1 Equilibrium urea denaturation of bovine and caprine βlg

Bovine and caprine βlg exhibit a similar level of resistance towards denaturation by urea, a chemical denaturant. This suggests that the amino acid substitutions between the bovine and caprine orthologues of βlg do not significantly affect their conformational stability.

To assess the level of denaturation occurring upon addition of urea requires a means of measuring the level of foldedness of the protein. In this case, the circular dichroism (CD) of βlg at 218 nm was monitored. When folded, there is a negative peak in the CD spectra at 218 nm (see figure 3.5) for both caprine and bovine βlg. This is characteristic of a β-sheet structure, which reflects the predominance of β-sheets in the native structure of these proteins. Following incubation in urea for 4 hours the magnitude of the dip at 218 nm is reduced, reflecting disruption of the secondary structure. This presents a useful feature with which to monitor βlg denaturation by urea.

Each of the three βlg variants exhibited similar urea denaturation behaviour (Figure 5.1). The profiles obtained indicate a two-state transition between a folded and an unfolded state. The data was collected for duplicate samples and the replicate data was averaged, to which a sigmoid curve was fitted. The midpoint of the sigmoid can be used as a parameter to compare the relative stability of each βlg protein. These were very similar at 4.8, 4.5 and 4.6 M for bovine βlg A, bovine βlg B and caprine βlg, respectively. These results suggest that bovine βlg A is marginally more stable to urea denaturation than bovine βlg B and caprine βlg, while the latter two proteins share a similar resistance to unfolding by urea.

To assess whether equilibrium (between the folded and unfolded states of βlg) has been reached within the incubation time of four hours, and also to assess the unfolding reversibility of βlg, samples of urea-denatured protein were diluted and their spectra re-measured. As can be seen in Figure 5.2, protein refolding following urea denaturation is possible. A protein sample in the presence of 9 M urea for 4 hours has a spectrum completely distinct from that of protein at 1 M urea. When the urea concentration is diluted from 9 M to 1 M the spectrum returns to that of the protein at 1 M urea. While some protein unfolding will have already

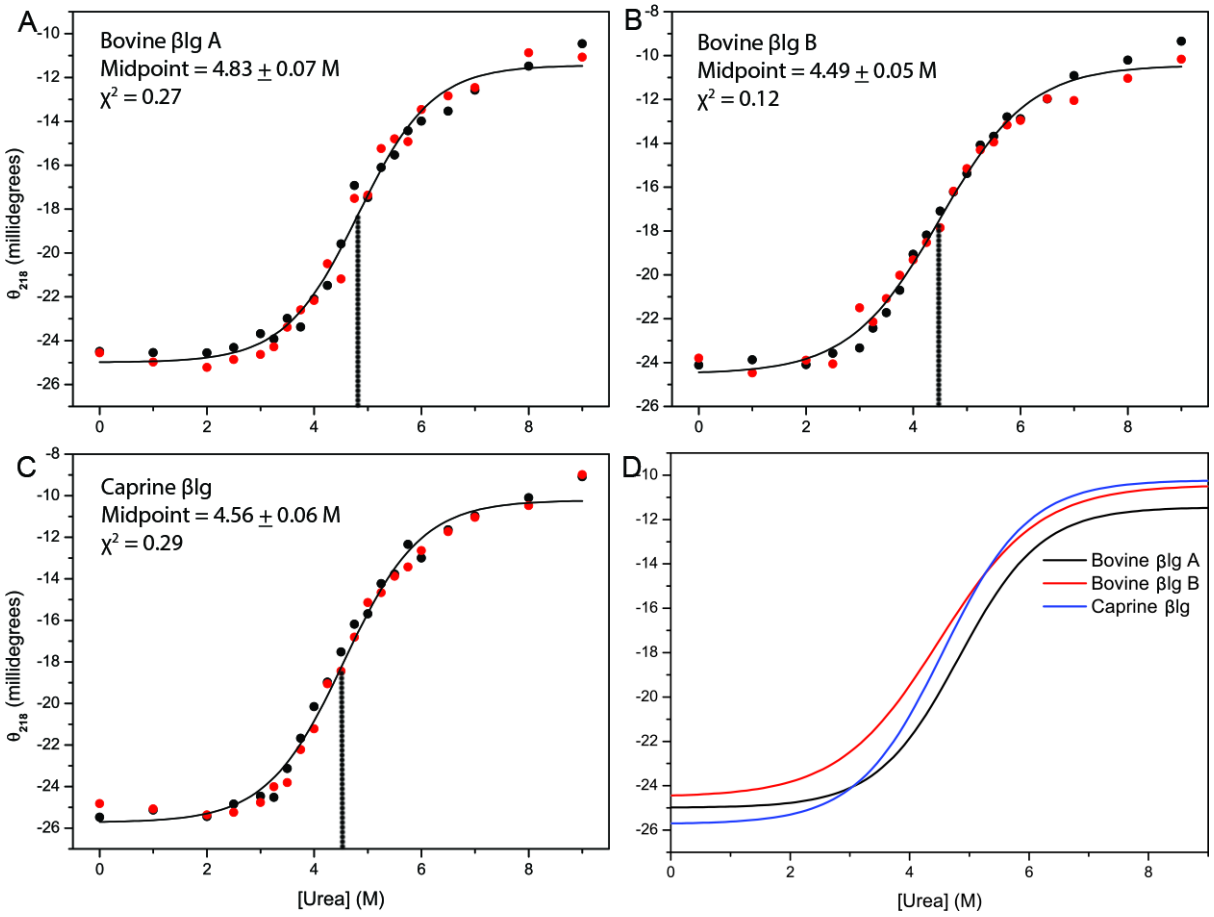


Figure 5.1: Urea denaturation curves of A) Bovine βlg A, B) Bovine βlg B and C) Caprine βlg. Replicate data (circles) was averaged and a sigmoid curve was fitted to the data (line). D) Overlaid urea denaturation curves (Bovine βlg A: black line, bovine βlg B: red line and caprine βlg: blue line).

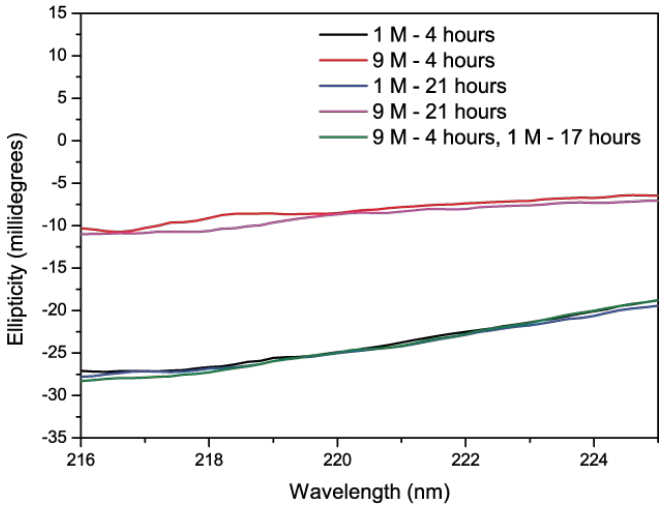


Figure 5.2: The circular dichroism of samples at 1 M (4 hours: black line, 21 hours: blue line) and 9 M urea (4 hours: red line, 21 hours: magenta line), and diluted from 9 M to 1 M urea (green line).

occurred at 1 M urea, it is assumed that if the urea was completely removed, for instance by dialysis, the protein would be able to completely refold. It is also apparent that four hours is sufficient time to reach equilibrium as the spectrum of the sample of protein in 9 M urea for 4 hours is very close to that at 9 M urea for 21 hours.

5.2.1.2 Thermal denaturation of bovine and caprine βlg

The thermal stability of caprine βlg displays a dependence on pH that differs from that of bovine βlg. Under conditions where the free thiol group, usually buried in native structures of βlg, is available to form heat-induced disulfide bonds between subunits, caprine βlg visibly precipitates when heated at pH values below 6.9. No precipitation of this sort was seen for bovine βlg at the pH values examined.

5.2.1.2.1 Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) is a method that monitors the unfolding of proteins as they are heated at a steady rate. The method utilises a fluorescent dye, such as SYPRO® Orange Protein Gel Stain, that only begins to fluoresce when it encounters the hydrophobic sites of proteins exposed upon unfolding. The midpoint of the unfolding transition is a useful parameter which can be used to compare the relative stability of proteins to thermal denaturation.

DSF was employed to compare the thermal stability of bovine and caprine βlg. Unfortunately SYPRO® Orange is not a suitable reporter for βlg unfolding (Figure 5.3). Rather than the desired sharp increase in fluorescence, the fluorescence increases steadily with a high level of background fluorescence. This is likely due to the fact that βlg proteins contain a hydrophobic pocket, the calyx, which presumably binds to the dye and causes a high background of fluorescence that masks any unfolding transition.

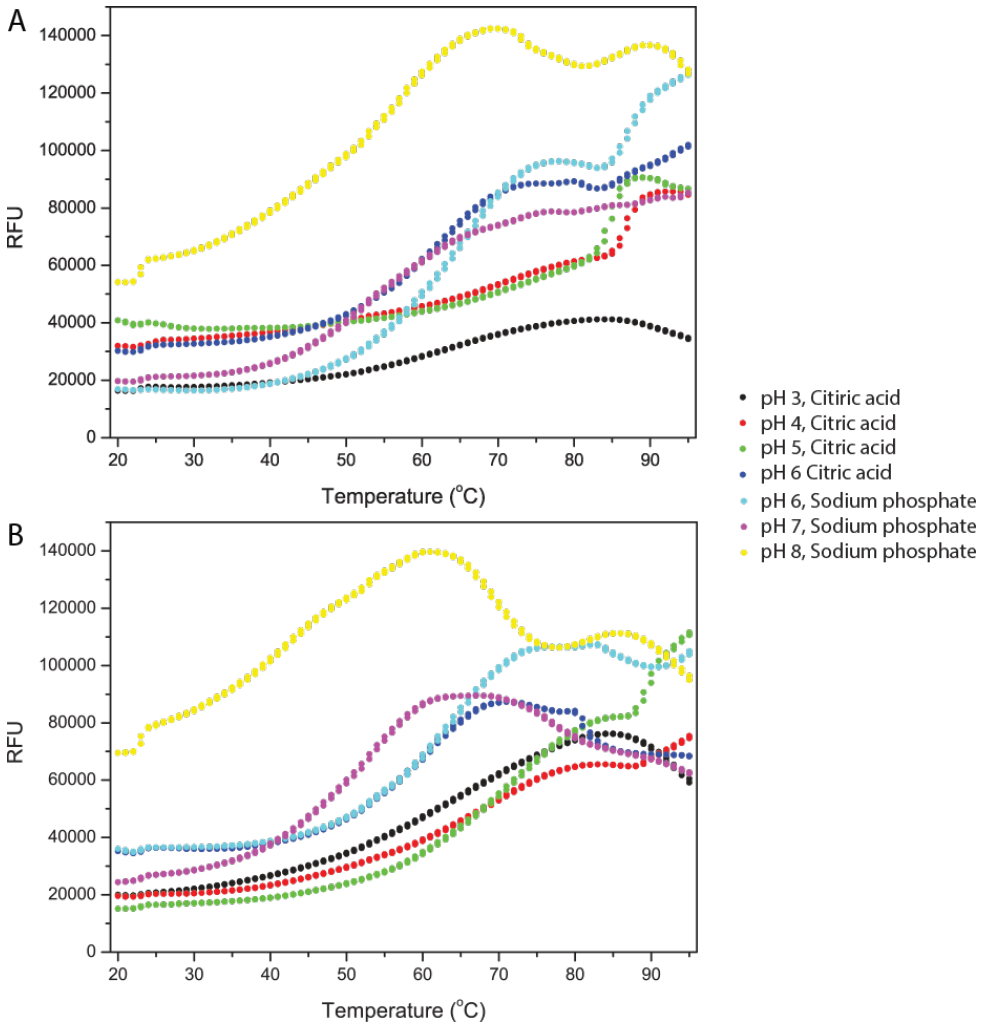


Figure 5.3: Bovine βlg A and caprine βlg were examined at six different pH values: pH 3, 4, 5 and 6 (0.1 M citric acid, 0.2 M sodium phosphate) and pH 6, 7 and 8 (0.2 M sodium phosphate) in the presence of SYPRO® Orange dye.

An alternative dye to SYPRO® Orange was sought. The dye 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM) gave well defined transitions upon thermal treatment of βlg. CPM evolves fluorescence when it reacts with and covalently binds to thiol groups (Alexandrov *et al.*, 2008) , while it is non-fluorescent in its unbound state. It can thus be used as a reporter for protein unfolding if the protein contains a buried thiol that is exposed upon unfolding.

βLg proteins contain five cysteine residues in each monomer. Four of these are involved in intermolecular disulfide bonds, while the fifth (Cys121) is buried underneath the α-helix, which is packed against the central β-barrel (see figure 4.9). It is known that this thiol group can become exposed during thermal denaturation (Manderson *et al.*, 1998) and that it can then

participate in disulfide exchange reactions with other βlg subunits, forming protein aggregates. Therefore, it follows that CPM would be useful to monitor structural changes that occur as βlg is heated, as the CPM dye would react with Cys121 as it becomes exposed.

All three orthologues of βlg share similar melting temperatures with respect to each other at each pH examined. Bovine βlg A and B and caprine βlg were analysed at four pH values: 6, 6.5, 7 and 8 (in 50 mM sodium phosphate, 50 mM NaCl). Unfortunately, CPM cannot be used outside of this range, as at pH values lower than 6 the rate of CPM-thiol adduct formation decreases, while at pH values above 8 the selectivity of the dye is reduced and CPM may start to react with primary amines (Alexandrov *et al*., 2008) . As the fluorescence signal fluctuates substantially at high temperatures, the top plateau of the unfolding profile was held constant at the maximum fluorescence obtained following the melting transition, as in Alexandrov *et al*. (2008) (Figure 5.4). A Boltzmann sigmoid curve was then fitted to the data in Origin. The midpoint of this curve was designated as the melting temperature (T_m).

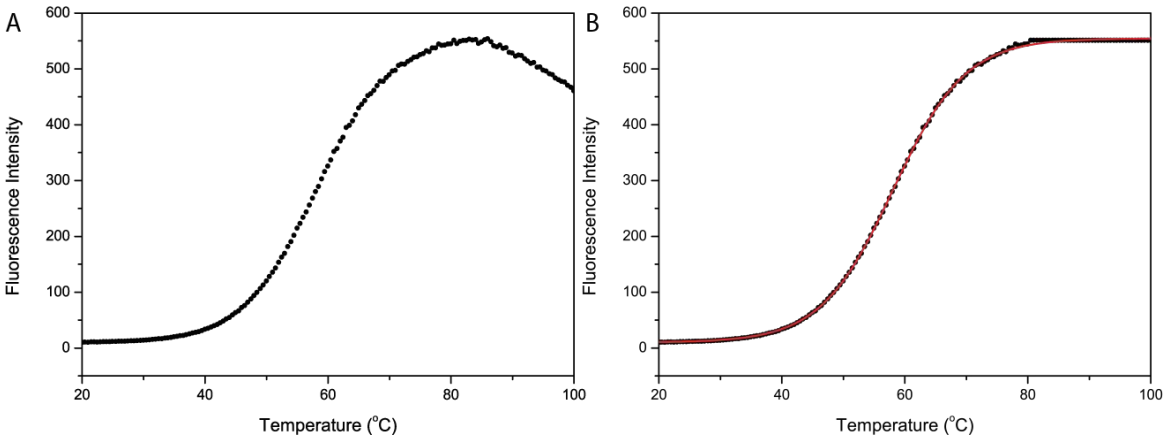


Figure 5.4: A representative melting curve obtained for bovine βlg A at pH 7 in the presence of CPM. The protein was heated from 20 to 100°C, at 1°C per minute. A) Initial data obtained, B) The plateau of the unfolding profile was held constant at the maximum fluorescence obtained following the melting transition. Data (black circles) are overlaid with a sigmoid curve (red line).

The thermal stability of each protein appears to decrease considerably from pH 6 to pH 8 (Figure 5.5). However, it should be kept in mind that the evolution of fluorescence is dependent

on the rate of the reaction between CPM and the free thiol group of βlg. This will be affected by, not only the availability of the thiol due to heat-induced structural changes, but also the reactivity of CPM. The shortcoming of this technique for quantifying the thermal stability of proteins as a function of pH is that the reactivity of CPM is strongly dependent on pH. While it has been shown that the thermal stability of bovine βlg increases with decreasing pH (Kella and Kinsella, 1988) , it is difficult to distinguish this, due to the fact that CPM-thiol adduct formation is reduced at lower pH values, and may therefore take longer to react with thiol groups as they become exposed. It is also unlikely that such a large decrease in stability occurs from pH 6 to pH 8 (40°C). This is more likely a symptom of the reduced specificity of CPM at pH values above 8 which will increase the number of targets that CPM may react with. Thus, comparisons can only be made between proteins at a single pH value. What these data serve to show is that each of the βlg orthologues have a similar reaction to thermal treatment at each pH under these conditions.

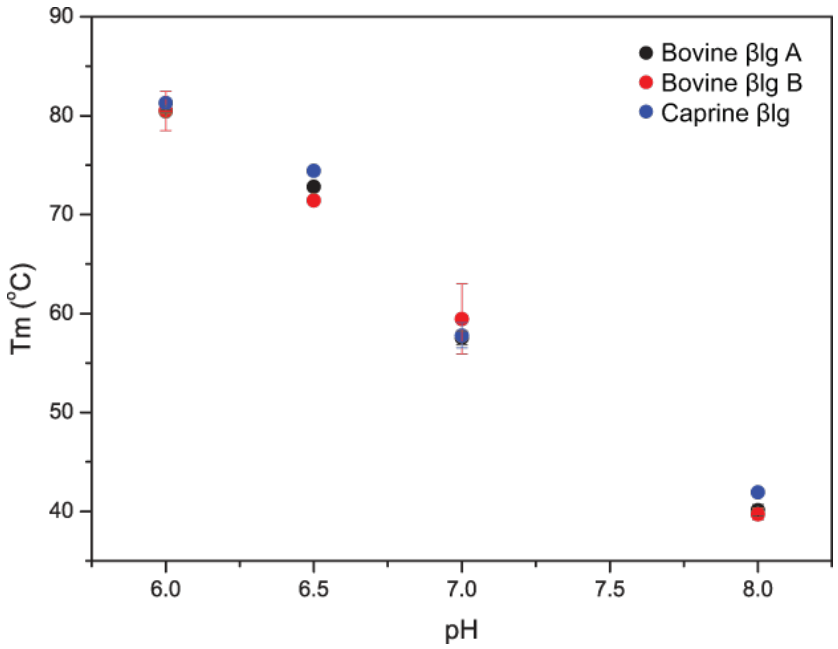


Figure 5.5: Melting temperatures obtained by differential scanning fluorimetry. Proteins were heated in the presence of CPM from 20 to 100°C at 1°C per minute. Values shown are averages of the melting temperatures obtained from duplicate data sets. Error bars indicate the standard deviation from the mean.

5.2.1.2.2 Circular dichroism spectroscopy

The thermal stability of bovine and caprine βlg as a function of pH was examined by monitoring changes to the near-UV CD spectra upon thermal treatment. Features of near-UV CD spectra arise due to the chemical environment surrounding aromatic residues that absorb in the near-UV range. If the unfolding of proteins alters these environments, this will be reflected in changes to the near-UV CD spectra. Figure 5.6 shows the near-UV CD spectrum of bovine βlg before and after heating. The differences in ellipticity at 286 nm and 293 nm are most likely consequences of changes to the environment of residue Trp19, reflecting structural changes that occur within the calyx of βlg as a result of heat treatment (Manderson *et al.*, 1999) . The largest difference is observed at 293 nm, and thus this wavelength was chosen to monitor the unfolding of βlg during heating. It is worth noting that the second spectrum (red line) was obtained after the temperature had returned to 25°C, indicating that the thermal-induced structural changes are not reversible.

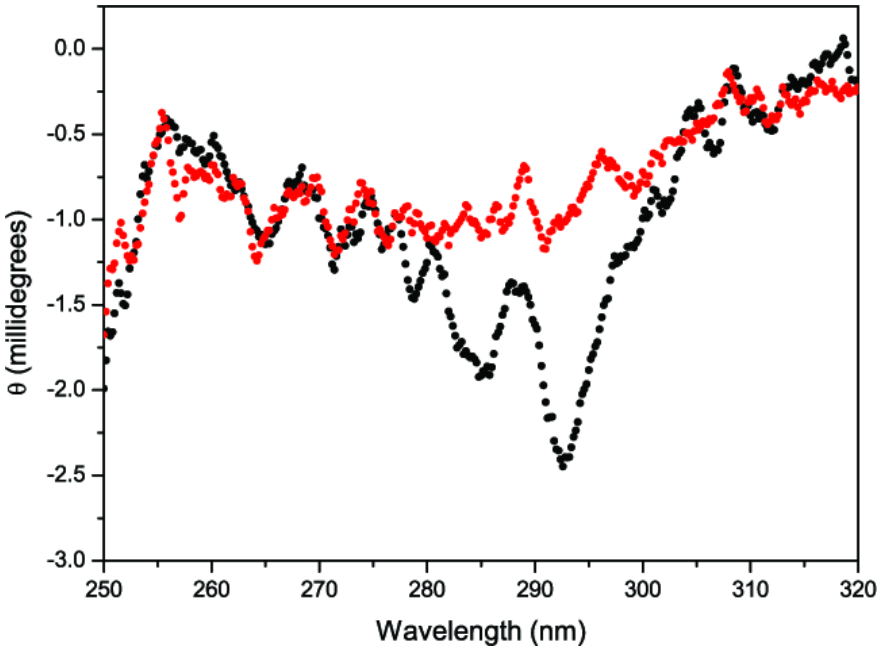


Figure 5.6: Changes in the near-UV CD spectra of bovine βlg A before (black line) and after (red line) heating to 95°C.

Bovine and caprine βlg were analysed at seven pH values: 3.5 (0.02 M citric acid/sodium citrate, 0.1 M NaCl), 6.5, 6.7, 6.9, 7.1, 7.3 and 7.4 (0.02 M sodium phosphate, 0.1 M NaCl). Each sample

was heated in a CD spectrometer, while being stirred with a magnetic stir bar, from 25°C to 95°C (see section 2.6.2.1.1). At each degree the ellipticity at 293 nm was recorded. A Boltzmann sigmoid curve was then fitted to the data in Origin (see Figure 5.7 for a representative denaturation profile). The midpoint of the curve was designated as the melting temperature (T_m). Data were collected in triplicate for each pH value.

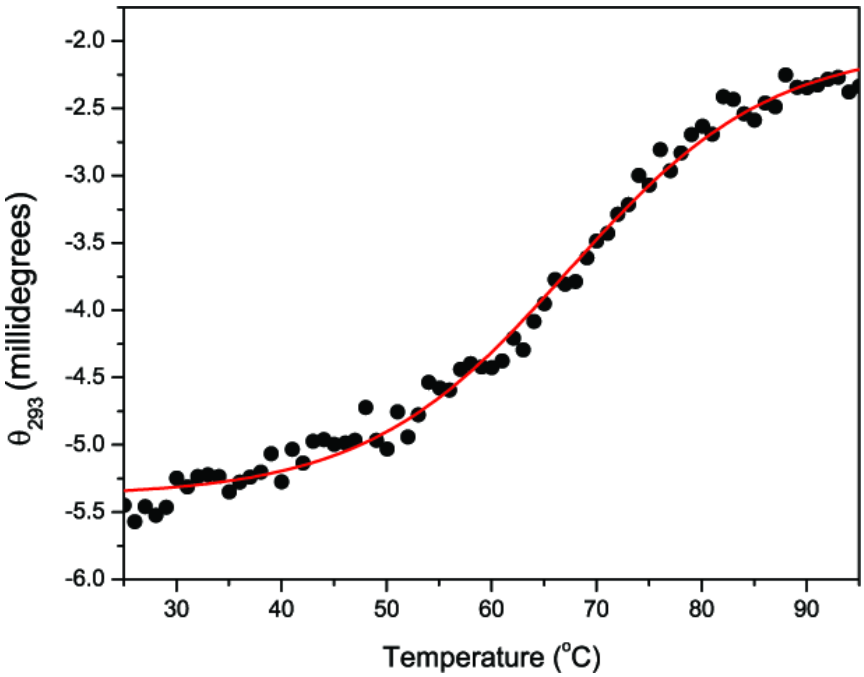


Figure 5.7: A representative melting curve obtained for bovine βlg A at pH 7.4. Protein was heated from 25 – 95°C. Data (black circles) are overlaid with a Boltzmann sigmoid curve (black line).

Manderson, Creamer and Hardman (Manderson *et al.*, 1999) obtained similar shaped thermal denaturation profiles and midpoint temperatures (Figure 5.8 and Table 5.1) for the bovine variants of βlg, also by monitoring the near-UV CD intensity at 293 nm. A significant difference between their protocol and that utilised here is that their measurements were made following a heating period of 10 minutes at a certain temperature (40 – 94°C) and a period of cooling on ice. Therefore, they were effectively monitoring the irreversible changes that occur upon heating, while here the structural changes that occur during heating are being observed. The fact that these findings are very similar indicates that the majority of the structural changes that occur in βlg during heating are irreversible.

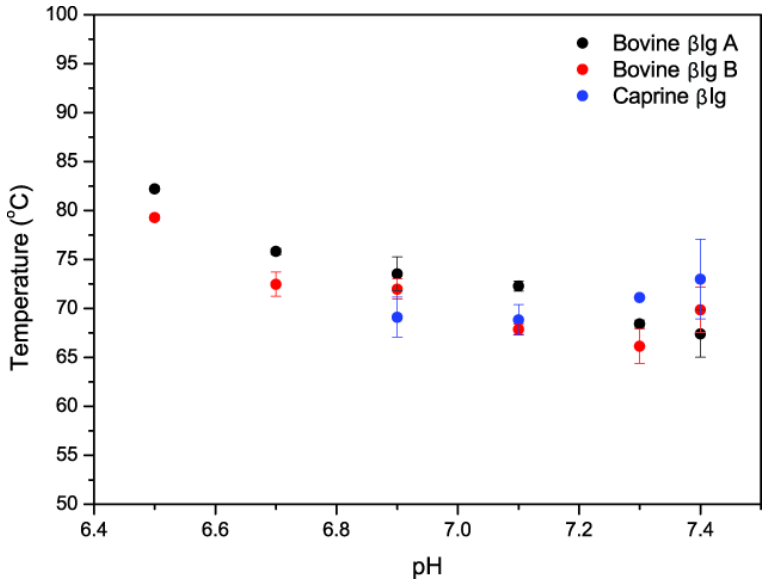


Figure 5.8: Midpoints of thermal denaturation of bovine and caprine βlg proteins. The T_m values shown are the averages of triplicate data sets, with the standard deviation shown as error bars.

Table 5.1: Midpoint temperatures for thermal denaturation of βlg proteins derived from the change in ellipticity at 293 nm.

pH	Bovine βlg A	Bovine βlg A[1]	Bovine βlg B	Bovine βlg B[1]	Caprine βlg
6.5	82.2 ± 0.3		79.3 ± 0.3		-
6.7	75.8 ± 0.3	74.7 ±0.3	72.5 ± 1.2	72.8 ± 0.0	-
6.9	73.6 ± 1.7		72.0 ± 1.0		69.1 ± 2.1
7.1	72.3 ± 0.5		67.9 ± 0.5		68.9 ± 1.5
7.3	68.5 ± 0.4		66.1 ± 1.7		71.1 ± 0.0
7.4	67.4 ±2.4	67.0 ± 0.2	69.9 ± 2.3	66.2 ± 0.1	73.0 ± 4.1

[1] Values from (Manderson *et al.*, 1999) .

Unfortunately, this technique is not suitable for deriving a T_m for bovine or caprine βlg proteins at pH 3.5 as they do not fully unfold by the time the CD spectrometer reaches 95°C (the suggested temperature limit of the machine). Therefore a sigmoid curve cannot be fitted to the data. Needless to say, each of these proteins exhibited similar, pronounced, thermal stability at pH 3.5 (Figure 5.9). Kella and Kinsella (1988) also observed an increase in the stability of bovine βlg proteins with decreasing pH. UV difference spectra of βlg as a function of pH suggest that this may be due to a reduction in the conformational flexibility of βlg with decreasing pH (Kella and Kinsella, 1988) .

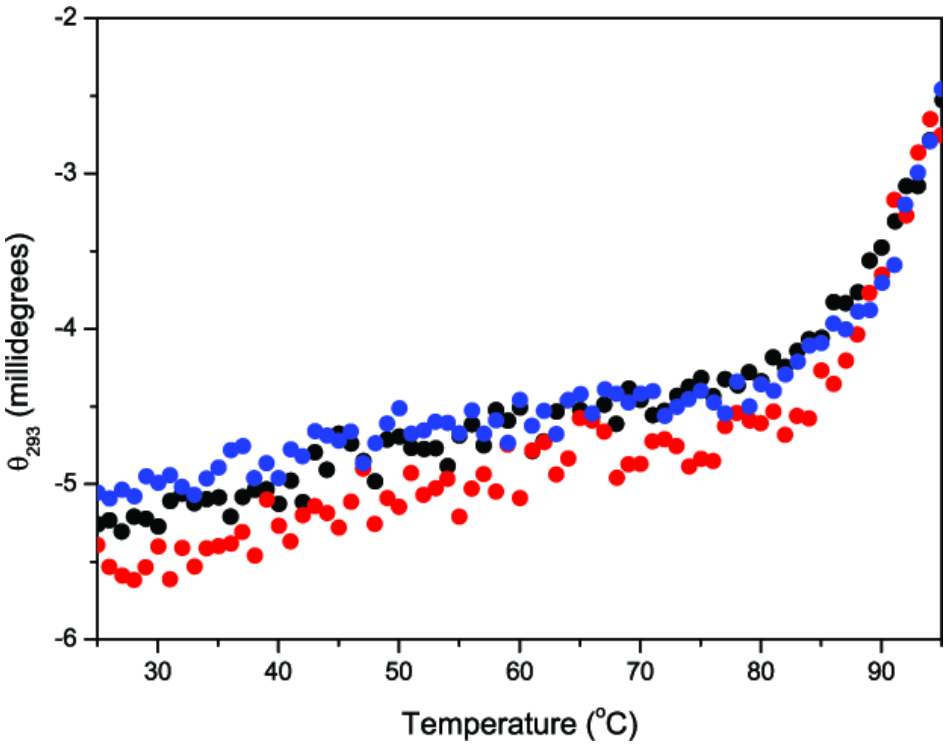


Figure 5.9: Thermal denaturation profiles of bovine βlg A (black), bovine βlg B (red) and caprine βlg (blue) at pH 3.5.

Caprine βlg shows a different trend in thermal stability as a function of pH than bovine βlg A and B (Figure 5.8). The bovine A and B variants of βlg show an increase in the T_m from pH 7.4 to pH 6.5 (67.4°C – 82.2°C and 69.9°C – 79.3°C, respectively). On the other hand, the T_m of caprine βlg decreases slightly from pH 7.4 to pH 6.9 (73°C – 69.1°C). The T_m for caprine βlg at pH values of 6.7 and 6.5 could not be quantified as it was not possible to fit a sigmoid curve to

the data obtained at these pH values. As can be seen in Figure 5.10, at around 80°C a sharp drop in the CD ellipticity occurs. This coincides with a sharp rise in the high tension voltage trace (Figure 5.10 B), suggesting that the absorbance of the sample has increased suddenly. When these samples were removed from the CD spectrometer there was noticeable precipitation within the cuvette (Figure 5.10 C). This was not observed for any of the bovine βlg A and B samples at any pH.

5.2.1.2.3 Analysis of thermal denaturation experiments

The precipitation seen following the heat treatment of caprine βlg at pH values below 6.9 is most likely due to a combination of disulfide-exchange induced aggregation and non-covalent interactions within the aggregates. βLg exists in solution as a dimer of two non-covalently linked subunits. Within each subunit there are five cysteine residues. Four of these form disulfide bonds (Cys66 – Cys160 and Cys109 – Cys119), while one remains as a free thiol (Cys121) that is buried in the folded, native structure. During heat denaturation of bovine βlg, the monomers dissociate and then begin to unfold (Roefs and Kruif, 1994) . This reveals the buried thiol, which can then react via disulfide-exchange with one of the disulfide bonds of another βlg monomer. This can occur recurrently, forming aggregates.

In the earlier DSF experiments (Section 5.2.1.2.1), CPM was used as a reporter of thermal-induced unfolding. This dye reacts covalently with the free thiol of βlg as it becomes exposed. If disulfide-exchange induced aggregation is required for the precipitation seen in the CD experiments, this explains why we did not see any precipitation of caprine βlg samples when conducting DSF with CPM. The DSF experiment using CPM also tells us that the availability of the thiol upon heating at pH 6.5 is similar for bovine and caprine βlg proteins (Figure 5.5). The midpoint for this transition is 73°C, 71°C and 74°C for bovine βlg A, bovine βlg B and caprine βlg, respectively. Thus, this raises the question: why do we not see any aggregation for bovine βlg at any of the pH values examined?

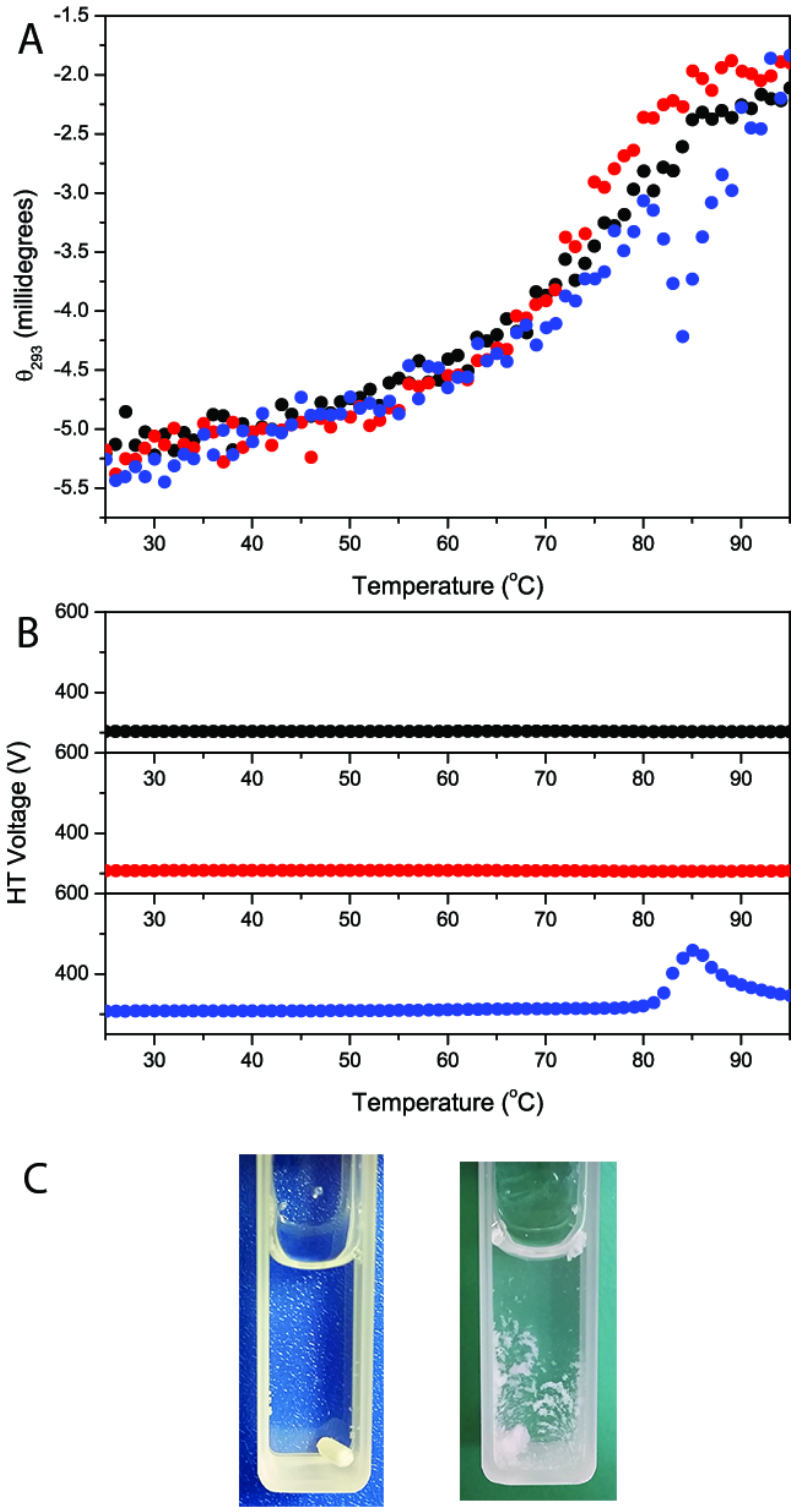


Figure 5.10: A) The change in ellipticity at 293 nm during thermal treatment of bovine βlg A (black line), bovine βlg B (red line) and caprine βlg (blue line). B) The high tension voltage trace for each βlg protein. C) Cuvettes following heat-treatment (bovine βlg A is on the left and caprine βlg is on the right). Note that the surface behind the cuvette of bovine βlg A is textured; this solution was transparent.

In 1998 Manderson, Hardman and Creamer carried out a thermal denaturation analysis of bovine βlg variants A, B and C (Manderson *et al*., 1998). They showed that heated solutions of bovine variants contain a different range of aggregated species, indicating differences in the behaviour between protein variants. In a non-reducing SDS-PAGE system, the heated protein samples separated into a series of bands with the most mobile band corresponding to the molecular weight of the monomer. At pH 6.7, variant B and C formed high molecular weight aggregates that were not able to enter the stacking part of the gel. This was not seen at pH 7.4, or in βlg A samples at either pH. It is believed that non-covalent associations also occur within aggregates, in addition to covalent disulfide-bonds, and that the net negative charge on the monomer is important in determining the size of the aggregates (Galani and Apenten, 1999; Manderson *et al*., 1998). At pH 6.7 the net negative charge on the βlg monomer increases in magnitude in the order bovine βlg C < bovine βlg B < bovine βlg A. The quantity of large aggregate material that was unable to enter the gel was greatest for bovine βlg C, then B, then A, which is the inverse order of net negative charge. As the pH is raised to 7.4 the net negative charge on the proteins increases, which may explain the increased proportion of protein able to enter the gel at this pH (Manderson *et al*., 1998).

Caprine βlg is estimated to carry a much lesser net negative charge at pH 7 than each of the bovine variant (see Table 4.1). This is due to the following amino acid substitutions in caprine βlg: D53N, D64G, D130K and E158G, that result in one extra positively charged amino acid and four less negatively charged amino acids compared to the bovine A and B βlg variants. Therefore, the pH values examined here are closer to the isoelectric point of caprine βlg than to those of the bovine variants. It may be that aggregates do form for each of the three βlg proteins but that those of caprine βlg form visible precipitates due to a greater amount of hydrophobically-driven, non-covalent associations within the aggregates. If the pH was reduced below pH 6.5, it is possible that we would start to see visible precipitation for the bovine βlg variants as well.

Interestingly, these results appear to follow the trend seen for the heat stability of milk. Goat milk heated at pH values below 6.9 suffers from significant precipitation (Anema and Stanley, 1998; Zadow *et al*., 1983), while in cow milk, if the pH is dropped below 6.4 these samples too begin to precipitate when heated (Zadow *et al*., 1983). This precipitation is largely due to the

association of denatured whey proteins, particularly βlg, with κ-casein. κ-casein is responsible for the stabilisation of casein micelles in milk by forming an interface between the hydrophobic caseins and the aqueous environment (Creamer *et al*., 1998). Dissociation of κ-casein (in complex with denatured whey proteins) leads to destabilisation of the casein micelles causing precipitation of the milk. Differences in the levels of association between βlg and κ-casein, and between the rates of dissociation of these complexes from casein micelles, is thought to be due to the higher calcium content of goat milk compared to cow milk (Montilla and Calvo, 1997; Wang *et al*., 2015; Zadow *et al*., 1983). However, it may be that this behaviour has more to do with the properties of βlg itself than previously thought.

5.2.2 Assessing the dynamic nature of βlg proteins

5.2.2.1 Molecular dynamics simulations

The amino acid substitutions between bovine and caprine βlg do not significantly alter the structures of these proteins or their conformational stability. There are, however, differences in the thermal denaturation and the self-association behaviours at certain pH values. It is possible that the dynamic nature of these proteins is altered with respect to each other.

While NMR spectroscopy has been used in the past to investigate the solution structure and dynamics of bovine βlg (Kuwata *et al*., 1999; Uhrínová *et al*., 2000), this is only possible at low pH (< 3), and low salt, where the protein is monomeric. These studies were key to understanding that the tertiary structure of bovine βlg at low pH is well maintained and very similar to that of a subunit within the dimer at neutral pH (Kuwata *et al*., 1999; Uhrínová *et al*., 2000). Given the interest in both the dynamic nature and the subunit association of these proteins, molecular dynamics simulations were used to compare the dynamic nature of bovine and caprine βlg. Molecular dynamics simulations have previously been applied to bovine βlg in order to understand the binding of ligands within the hydrophobic core and to understand the coupled effects of self-association and ligand binding (Bello *et al*., 2016, 2012; Evoli *et al*., 2014; Gutiérrez-Magdaleno *et al*., 2013; Yi and Wambo, 2015).

Here, the molecular dynamics of bovine βlg A and B and caprine βlg were simulated at 37°C, at pH 7, in 150 mM NaCl and in duplicate. 37°C was chosen to simulate physiological

temperatures, however βlg encounters a wide range of pH environments (from pH ~6.7 in milk, to as low as pH 3 in the stomach during digestion), thus the pH was kept at pH 7 for the comparison. The root mean square deviation (rmsd) between each structure sampled during the simulation and the initial structure can be plotted as a function of time, which allows visual assessment of whether each simulation has come to equilibrium. This can be carried out for the whole protein or for individual subunits of each protein to separate internal structural changes from the relative motions of subunits with respect to each other. While the individual protein subunits all come to equilibrium by 25 ns (Figure 5.11), rearrangements in subunit assembly result in the large rmsd values (> 0.5 nm) seen in some simulations (Figure 5.11 A, B, D and F).

5.2.2.2 Bovine and caprine βlg share the same regions of flexibility

In general, the regions that show the most flexibility are the same for subunits of bovine and caprine βlg. Using only trajectories from after the system has come to equilibrium (25 ns onwards), the root mean square fluctuations (rmsf) of each residue of each protein were assessed (Figure 5.12). This indicates which parts of each protein are the most dynamic. For bovine and caprine βlg, these regions are the loops found at each end of the calyx, along with the α-helix and the N and C termini (Figure 5.12). This is in agreement with NMR spectroscopy findings for bovine βlg (Kuwata *et al*., 1999; Uhrínová *et al*., 2000), and for previous simulations carried out on bovine βlg A and B (Bello *et al*., 2016; Evoli *et al*., 2014).

There is as much variation between simulation replicates as there are between chains of each βlg protein, l*et al*one between different orthologues of βlg, thus it is difficult to make any conclusions as to whether any of the orthologues exhibit greater flexibility relative to one another. When the fluctuations are averaged over each replicate and each protein chain it is apparent that the bovine and caprine orthologues of βlg share a similar degree of conformational flexibility. Overall, bovine βlg A is slightly more flexible than the other orthologues, however these differences are all within the error of the observations made (Figure 5.13). Others have demonstrated, via crystallographic B-factor analysis and Fourier transform infrared spectroscopy, that variant A of bovine βlg has greater structural mobility than variant B (Dong *et al*., 1996; Oliveira *et al*., 2001). Only the second replicate of bovine βlg

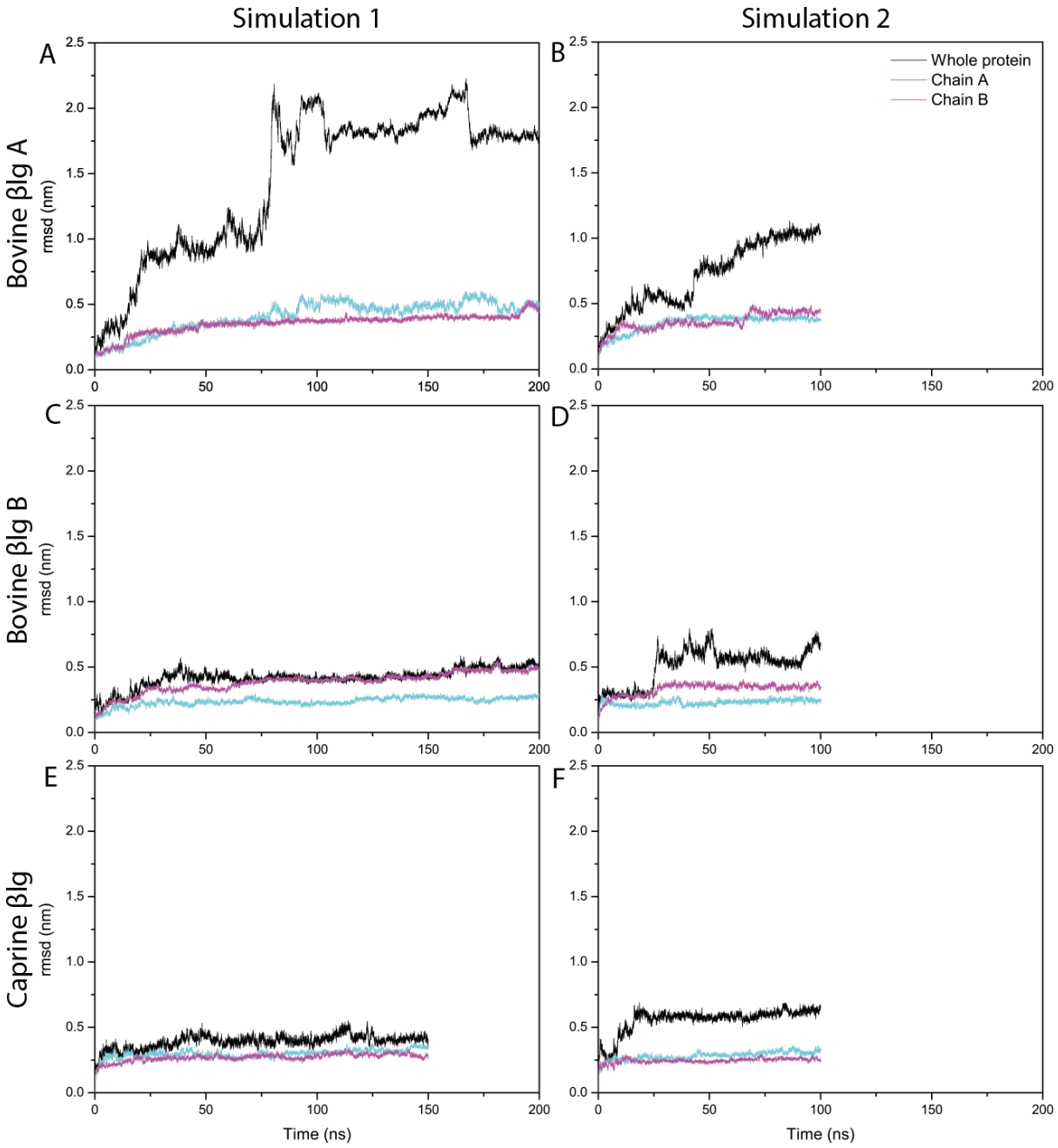


Figure 5.11: Cα atom positional rmsd as a function of time for the least squares fit of Cα atoms of each structure to the starting structure. Whole protein = black line, chain A = cyan line, chain B = magenta line. A and B) Bovine βlg A (PDB ID: 1BSO (Qin *et al.*, 1998b)), C and D) bovine βlg B (PDB ID: 1BSQ (Qin *et al.* 1999)), E and F) caprine βlg (PDB ID: 4TLJ (Crowther *et al.*, 2014)).

A was used in the rmsf comparison since, in the first replicate, bovine βlg A underwent dimer dissociation and it was thought that this may have influenced the enhanced conformational flexibility seen in the loops in this simulation (Figure 5.12 A).

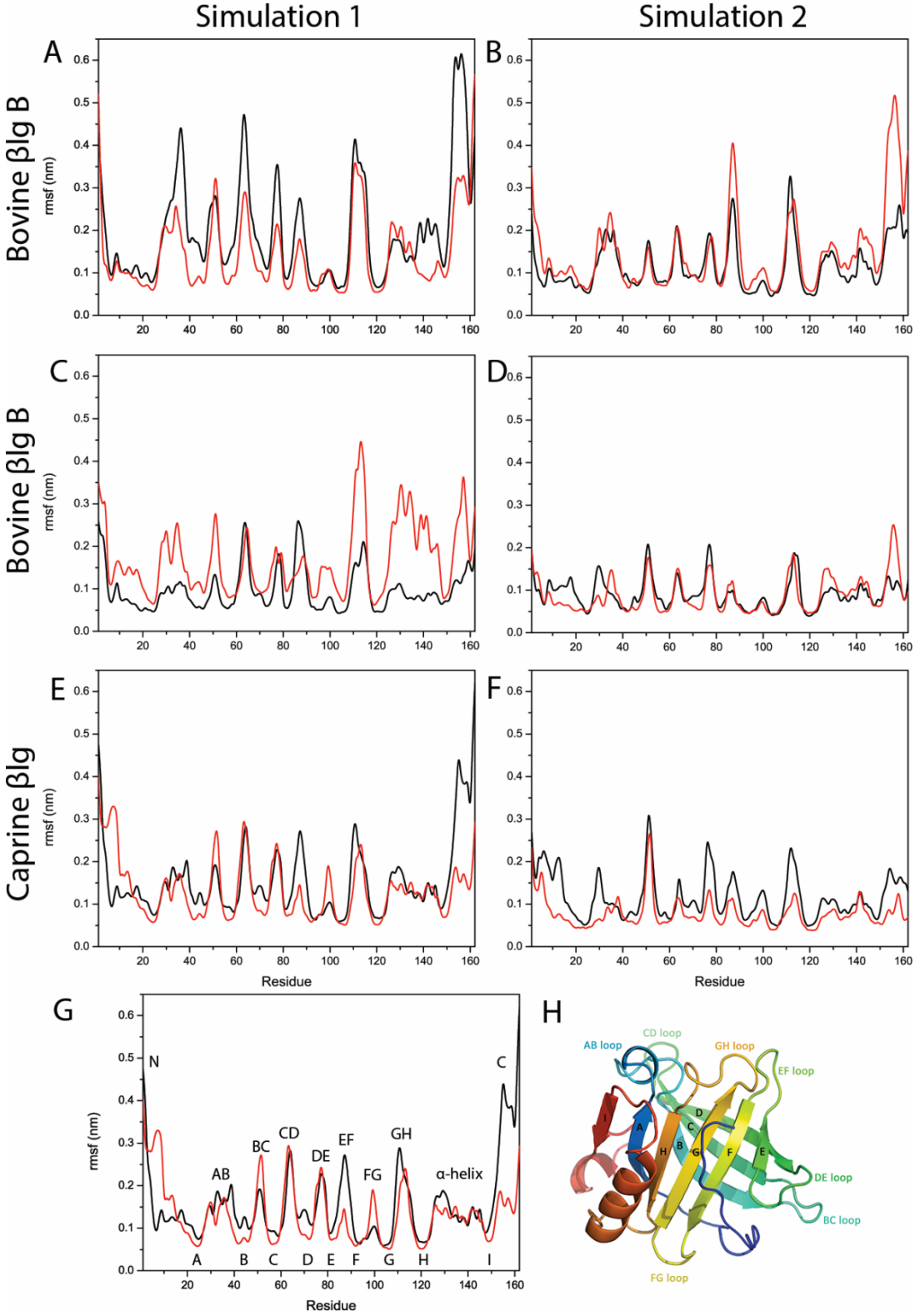


Figure 5.12: Root mean square fluctuations for residues of βlg structures for chain A (black) and chain B (red). A and B) bovine βlg A (PDB ID 1BSO), C and D) bovine βlg B (PDB ID 1BSQ), E and F) caprine βlg (PDB ID 4TLJ). G) Rmsf plot (E) with N and C termini, A-I strands and loops between strands labelled. H) Corresponding strands and loops are indicated on a cartoon structure of a single subunit of βlg coloured from the N terminus in blue to the C terminus in red.

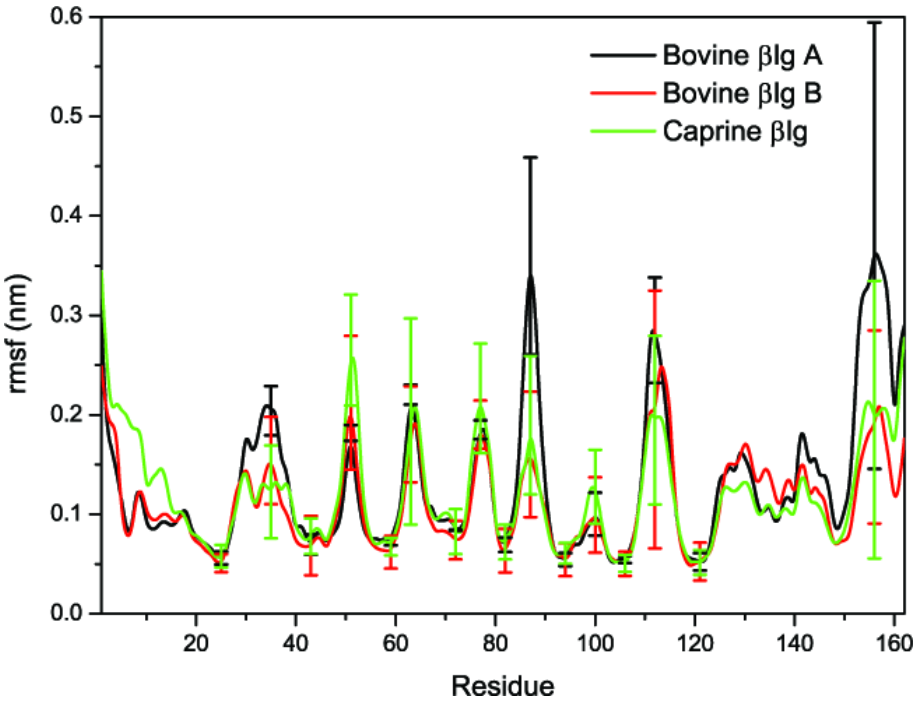


Figure 5.13: Average root mean square fluctuations of residues of bovine βlg A (black line), bovine βlg B (red line) and caprine βlg (green line). The standard deviation is shown as error bars for the peaks and troughs only, for visual clarity.

Based on structural mobility alone (Figure 5.13) it would seem that caprine βlg would exhibit a similar resistance towards proteolysis as to bovine βlg A and B. This suggests that the differing protein and fat composition of cow and goat milk, or changes to the protein structure during processing, make the caprine version of βlg more digestible. This remains to be supported by experimental data, such as a comparison of the *in vitro* digestion of these proteins in isolation.

5.2.2.3 There is flexibility at the dimer interface of βlg

Rearrangement and dissociation of subunits of bovine and caprine βlg can be observed when the course of the simulations are visualised in Visual Molecular Dynamics (VMD). The largest rearrangement of subunits is seen in the first simulation of bovine βlg A; at 75 ns the subunits completely separate (shown by the large jump in the rmsd value at 75 ns in Figure 5.11 A). The subunits tumble in solution briefly before coming back into contact at a site distant from the original dimer interface. When this simulation was repeated (i.e. the simulation process was

started again from the beginning using the same initial coordinates but different initial velocities) we do not see dimer dissociation, but the subunits do rotate considerably with respect to one another (shown by the gradual increase in rmsd in Figure 5.11 B).

Dimer dissociation was not observed in either of the replicate simulations of bovine βlg B. In the first simulation, the structure shows little deviation once equilibrium is reached, with only slight rotation observed at the dimer interface. In the duplicate simulation, however, a significant rotation of the subunits was observed at 25 ns (see Figure 5.11 D). There was a reasonable amount of movement between subunits for the remainder of this simulation.

Similarly for caprine βlg, the first simulation of caprine βlg (PDB ID 4TLJ) shows only slight movements at the dimer interface along the course of the simulation. In the duplicate simulation, a reasonably large rotation occurs early on in the simulation (~20 ns, see Figure 5.11 F) and the structure remains fairly stable from then on.

To ascertain whether the dimer dissociation seen for bovine βlg A was due to the starting structure chosen for simulation (PDB ID 1BSO), a different starting structure was selected (PDB ID 1BSY). Simulations were carried out using these new coordinates for bovine βlg A following the same process as for the first structure. During these simulations bovine βlg A again undergoes dimer dissociation (a large rotation of the subunits occurs at 75 ns and subunit dissociation occurs near the end of the simulation at ~90 ns, Figure 5.14). In the duplicate simulation, the subunits remain fairly stable relative to each other for the duration of the simulation, with only a small amount of rotation between subunits (Figure 5.14 B).

We know from AUC studies and by observing crystal structures that βlg dimers are only weakly held together by a small number of interactions (K_D values are within the range of 5 – 50 μM at neutral pH and 100 mM NaCl, see section 4.2.1 and figure 4.2). What is interesting, therefore, is that we only observe dimer dissociation in the case of bovine βlg A. This may be completely due to chance and had more replicate simulations been carried out for the other orthologues of βlg, we may have seen dimer dissociation occur for these proteins also. In milk, the concentration of these proteins will be much higher than analysed in AUC studies

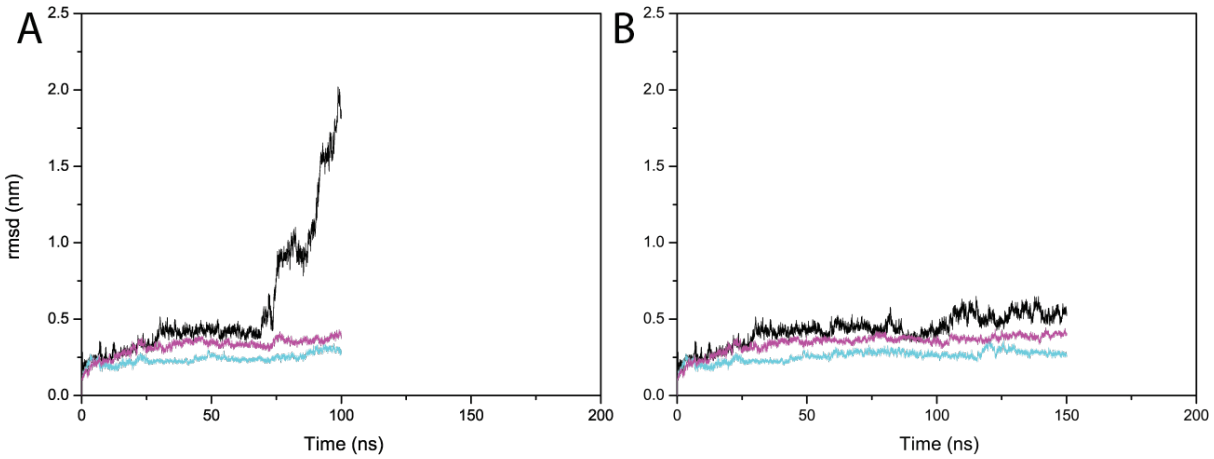


Figure 5.14: Cα atom positional rmsd as a function of time for the least squares fit of Cα atoms of each structure to the starting structure. A and B) Bovine βlg A (PDB ID: 1BSY): whole protein (black), chain A (cyan), chain B (magenta).

(~150 μM compared to the highest concentration observed of 55 μM). Milk is also a crowded solution, which will encourage self-association (Rivas *et al*., 2001).

Numerous interactions occur between the subunits of bovine and caprine βlg during these simulations that have not been observed in crystal structures of βlg orthologues. The crystallographic dimer interfaces of bovine and caprine βlg comprise interactions between the I β-strands and AB loops of each subunit (see figure 4.2). Four hydrogen bonds are regularly seen in crystal structures between His146 – Ser'150, Arg148 – Arg'148 (x2) and Ser150 – His'146 of the antiparallel I β-strands. Only in some structures of βlg can a network of hydrogen bonds be seen between the pairs of sidechain atoms of Asp33 and Arg40 and the main chain atoms of Ala34, located on the AB loops. In these structures, the charged sidechains of Asp33 and Arg40 also form a pair of salt-bridges. However, in other crystal structures, these sidechains (Asp33 and Arg40) are not always close enough to their symmetry-related counterparts to be classified as forming salt-bridges (< 0.4 nm (Xu *et al*., 1997)) (Gutiérrez-Magdaleno *et al*., 2013). While this was initially thought to be due to the pH of crystallisation, structures have since been solved that cover a range of pH values (such as PDB IDs 1BEB (Brownlow *et al*., 1997), and 1B8E (Oliveira *et al*., 2001) at pH 6.5 and 7.9, respectively) which do show these groups interacting at the dimer interface. These can be compared to other structures (such as PDB IDs 3BLG, 1BSY and 2BLG (Qin *et al*., 1998a) at pH 6.2, 7.1 and 8.2,

respectively) which all show broken salt-bridges. It is now thought that the presence of these salt-bridges is dependent on the crystal packing (Bello *et al.*, 2016). In the structures selected here as the starting coordinates for simulations, the Asp33 and Arg40 groups are only within 0.4 nm of one another in the structure of caprine βlg (PDB ID 4TLJ). In the structures of bovine βlg A and B, these groups are at least 0.6 nm from each other (Figure 5.15).

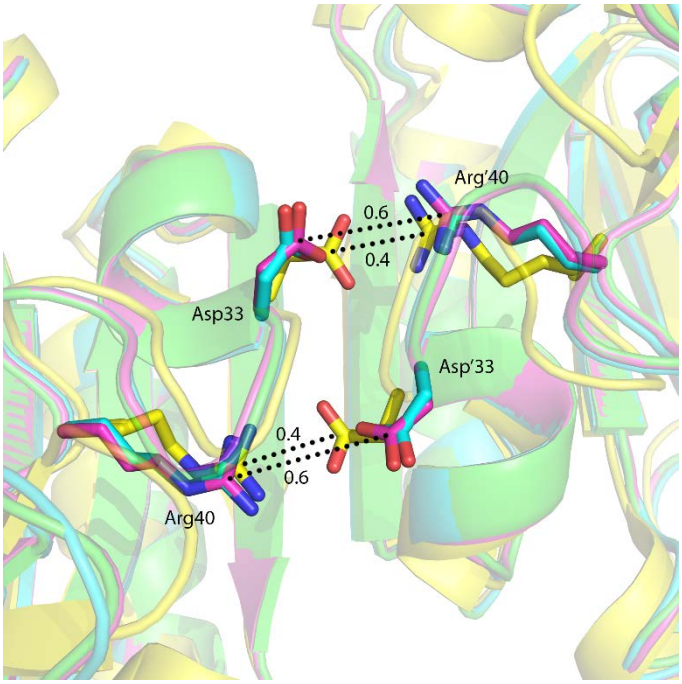
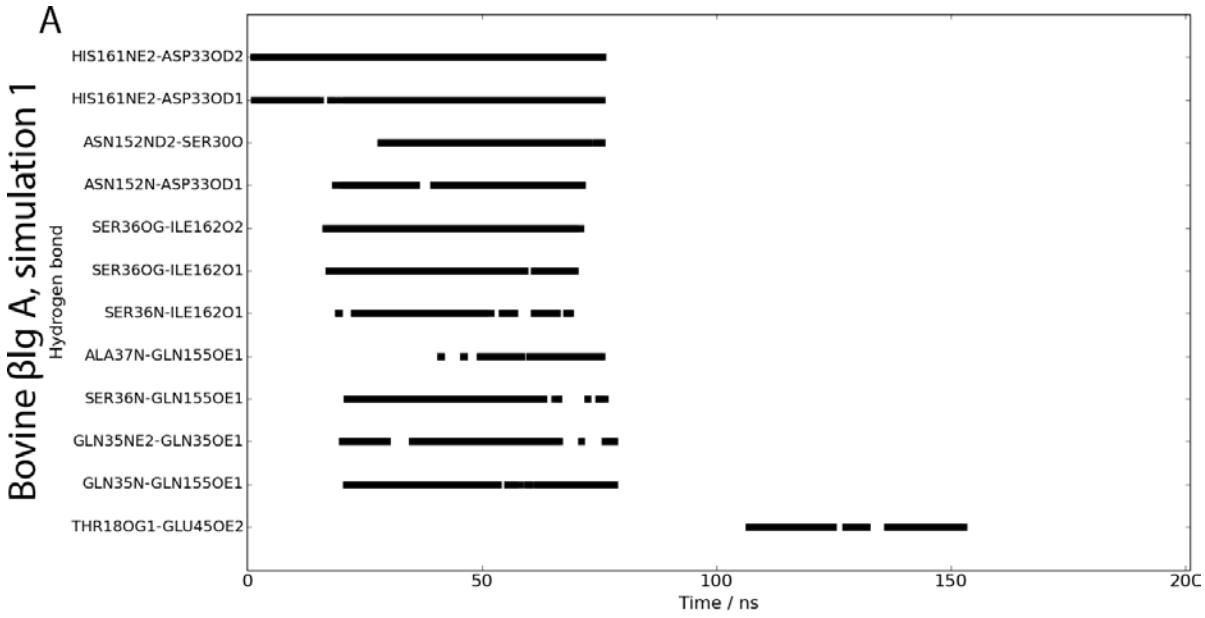
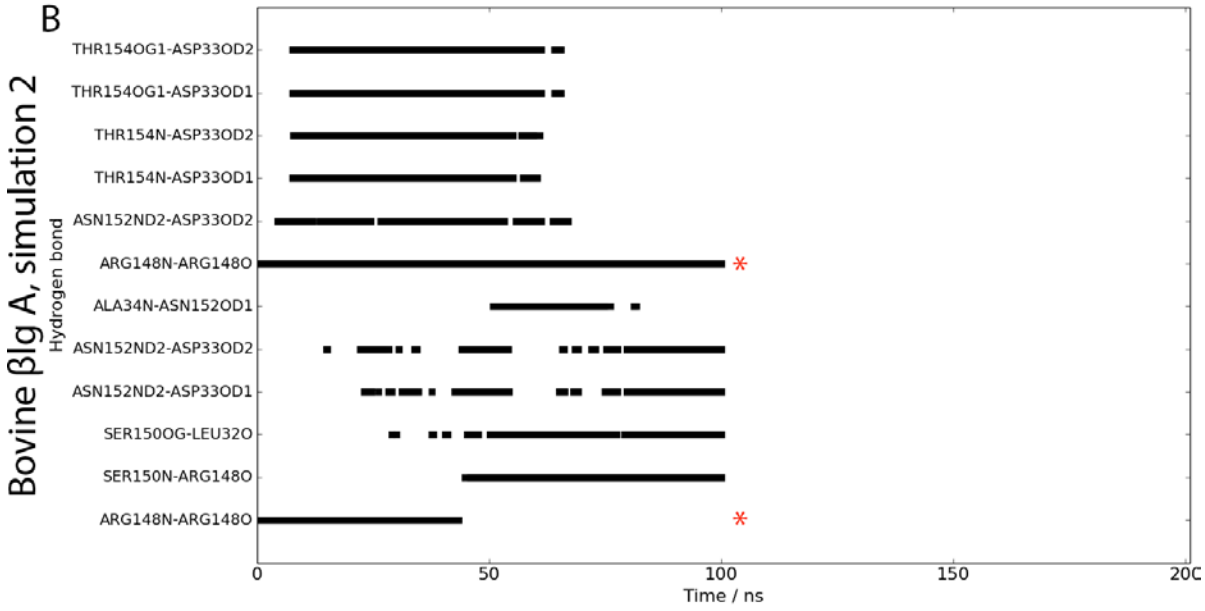
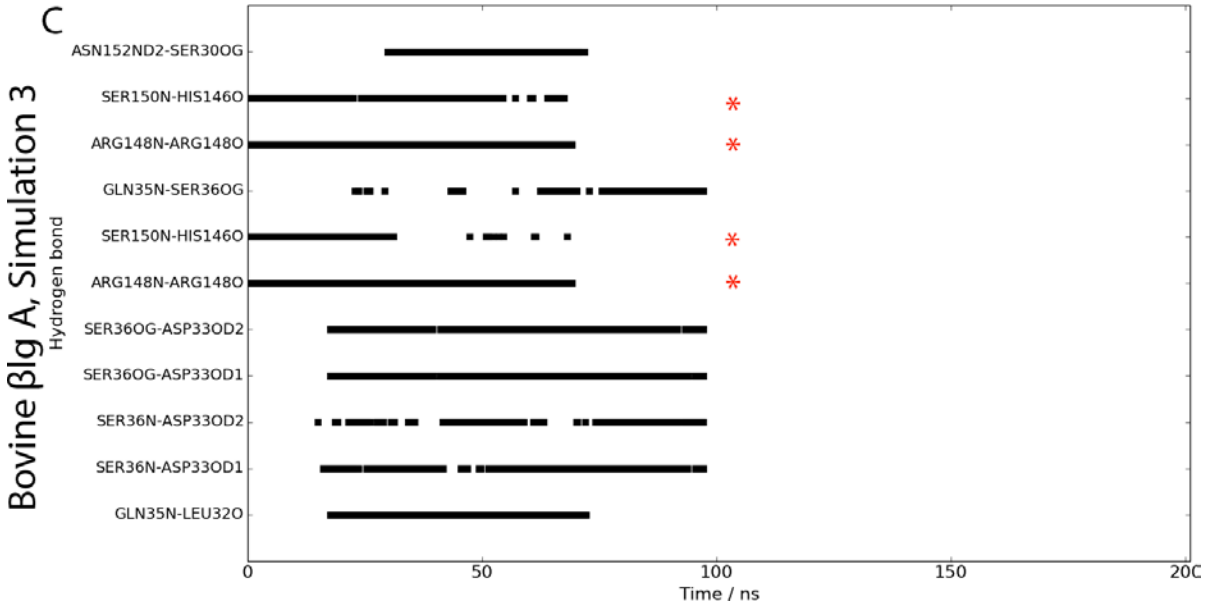


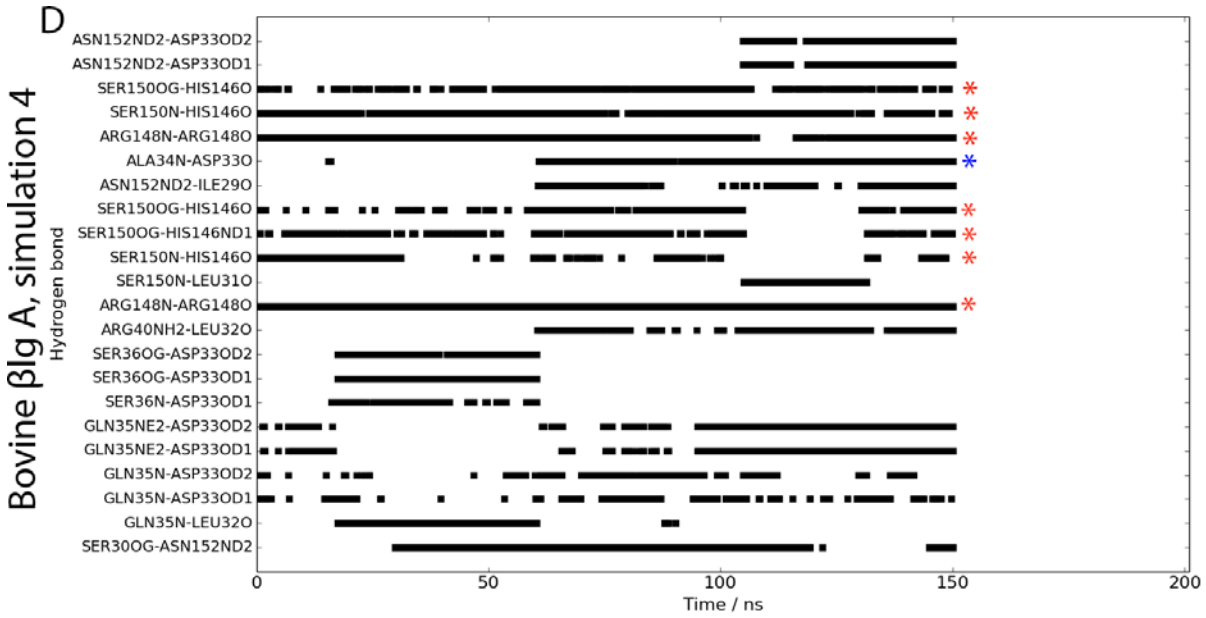
Figure 5.15: The distances between Asp33 and Arg40 groups of each subunit. Distances are indicated in nm. Bovine βlg A shown in green (PDB 1BSO) and magenta (PDB 1BSQ), bovine βlg B shown in cyan (PDB 1BSQ) and caprine βlg shown in yellow (PDB 4TLJ).

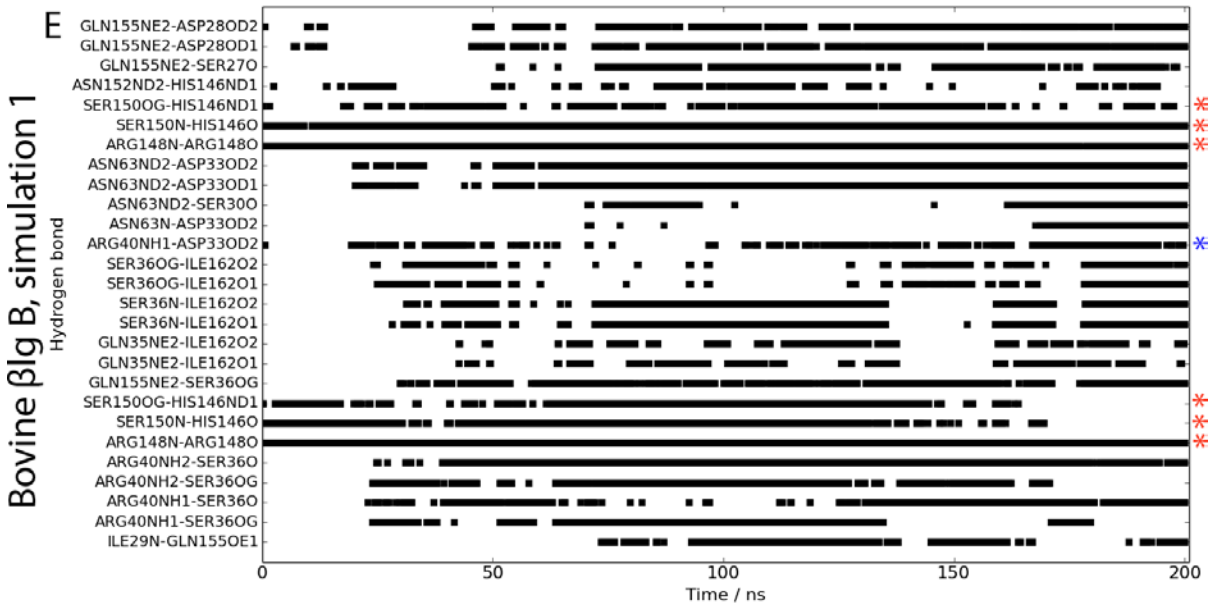
An analysis of the hydrogen bonds and salt-bridges that form between subunits during the course of the simulations reveals a difference in the network of interactions that occur between subunits to what is traditionally seen in crystal structures of βlg. With the exception of the first and third simulations of bovine βlg A (where dimer dissociation occurs), hydrogen bonds between His146, Ser150 (Ala150 in caprine structures) and Arg148 of the I β-strand are persistent throughout the course of the simulations (Figure 5.16, see the red asterisks). However, hydrogen bonding is not always seen between the Asp33, Ala34 and Arg40 residues

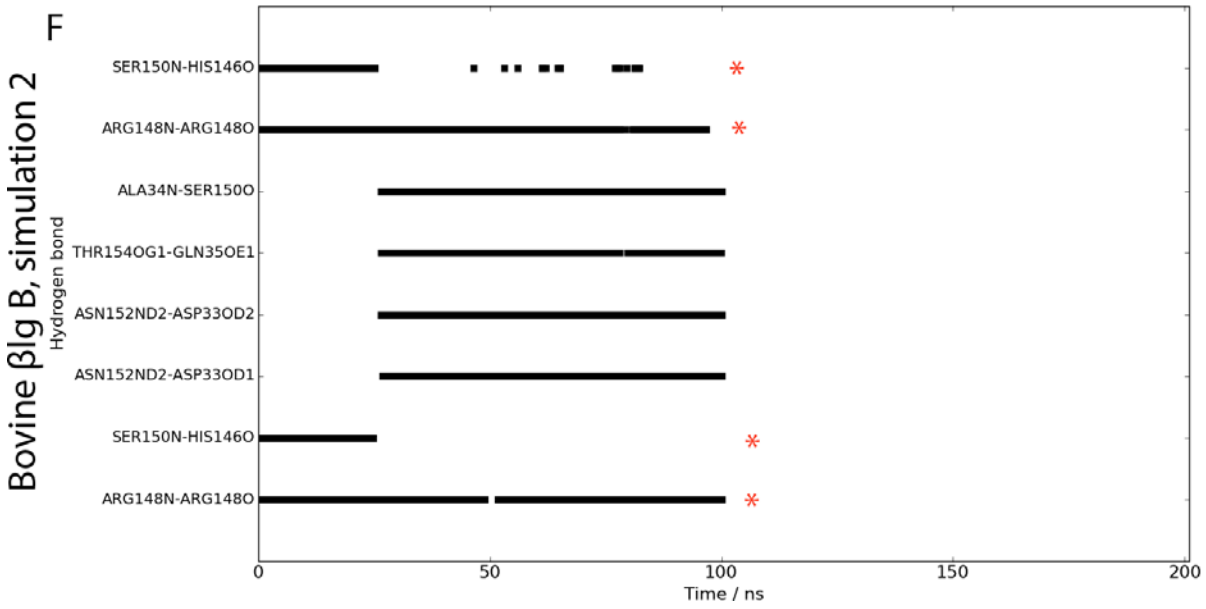

A
Bovine βlg A, simulation 1
Hydrogen bond
HIS161NE2-ASP33OD2
HIS161NE2-ASP33OD1
ASN152ND2-SER30O
ASN152N-ASP33OD1
SER36OG-ILE162O2
SER36OG-ILE162O1
SER36N-ILE162O1
ALA37N-GLN155OE1
SER36N-GLN155OE1
GLN35NE2-GLN35OE1
GLN35N-GLN155OE1
THR18OG1-GLU45OE2
0
50
100
150
200
Time / ns


B
Bovine βlg A, simulation 2
Hydrogen bond
THR154OG1-ASP33OD2
THR154OG1-ASP33OD1
THR154N-ASP33OD2
THR154N-ASP33OD1
ASN152ND2-ASP33OD2
ARG148N-ARG148O
ALA34N-ASN152OD1
ASN152ND2-ASP33OD2
ASN152ND2-ASP33OD1
SER150OG-LEU32O
SER150N-ARG148O
ARG148N-ARG148O
0
50
100
150
200
Time / ns


C
Bovine βlg A, Simulation 3
Hydrogen bond
ASN152ND2-SER30OG
SER150N-HIS146O
ARG148N-ARG148O
GLN35N-SER36OG
SER150N-HIS146O
ARG148N-ARG148O
SER36OG-ASP33OD2
SER36OG-ASP33OD1
SER36N-ASP33OD2
SER36N-ASP33OD1
GLN35N-LEU32O
0
50
100
150
200
Time / ns


D
Bovine βlg A, simulation 4
Hydrogen bond
ASN152ND2-ASP33OD2
ASN152ND2-ASP33OD1
SER150OG-HIS146O
SER150N-HIS146O
ARG148N-ARG148O
ALA34N-ASP33O
ASN152ND2-ILE29O
SER150OG-HIS146O
SER150OG-HIS146ND1
SER150N-HIS146O
SER150N-LEU31O
ARG148N-ARG148O
ARG40NH2-LEU32O
SER36OG-ASP33OD2
SER36OG-ASP33OD1
SER36N-ASP33OD1
GLN35NE2-ASP33OD2
GLN35NE2-ASP33OD1
GLN35N-ASP33OD2
GLN35N-ASP33OD1
GLN35N-LEU32O
SER30OG-ASN152ND2
0
50
100
150
200
Time / ns


E
Bovine βlg B, simulation 1
Hydrogen bond
GLN155NE2-ASP28OD2
GLN155NE2-ASP28OD1
GLN155NE2-SER27O
ASN152ND2-HIS146ND1
SER150OG-HIS146ND1
SER150N-HIS146O
ARG148N-ARG148O
ASN63ND2-ASP33OD2
ASN63ND2-ASP33OD1
ASN63ND2-SER30O
ASN63N-ASP33OD2
ARG40NH1-ASP33OD2
SER36OG-ILE162O2
SER36OG-ILE162O1
SER36N-ILE162O2
SER36N-ILE162O1
GLN35NE2-ILE162O2
GLN35NE2-ILE162O1
GLN155NE2-SER36OG
SER150OG-HIS146ND1
SER150N-HIS146O
ARG148N-ARG148O
ARG40NH2-SER36O
ARG40NH2-SER36OG
ARG40NH1-SER36O
ARG40NH1-SER36OG
ILE29N-GLN155OE1
0
50
100
150
200
Time / ns


F
Bovine βlg B, simulation 2
Hydrogen bond
SER150N-HIS146O
ARG148N-ARG148O
ALA34N-SER150O
THR154OG1-GLN35OE1
ASN152ND2-ASP33OD2
ASN152ND2-ASP33OD1
SER150N-HIS146O
ARG148N-ARG148O
0
50
100
150
200
Time / ns

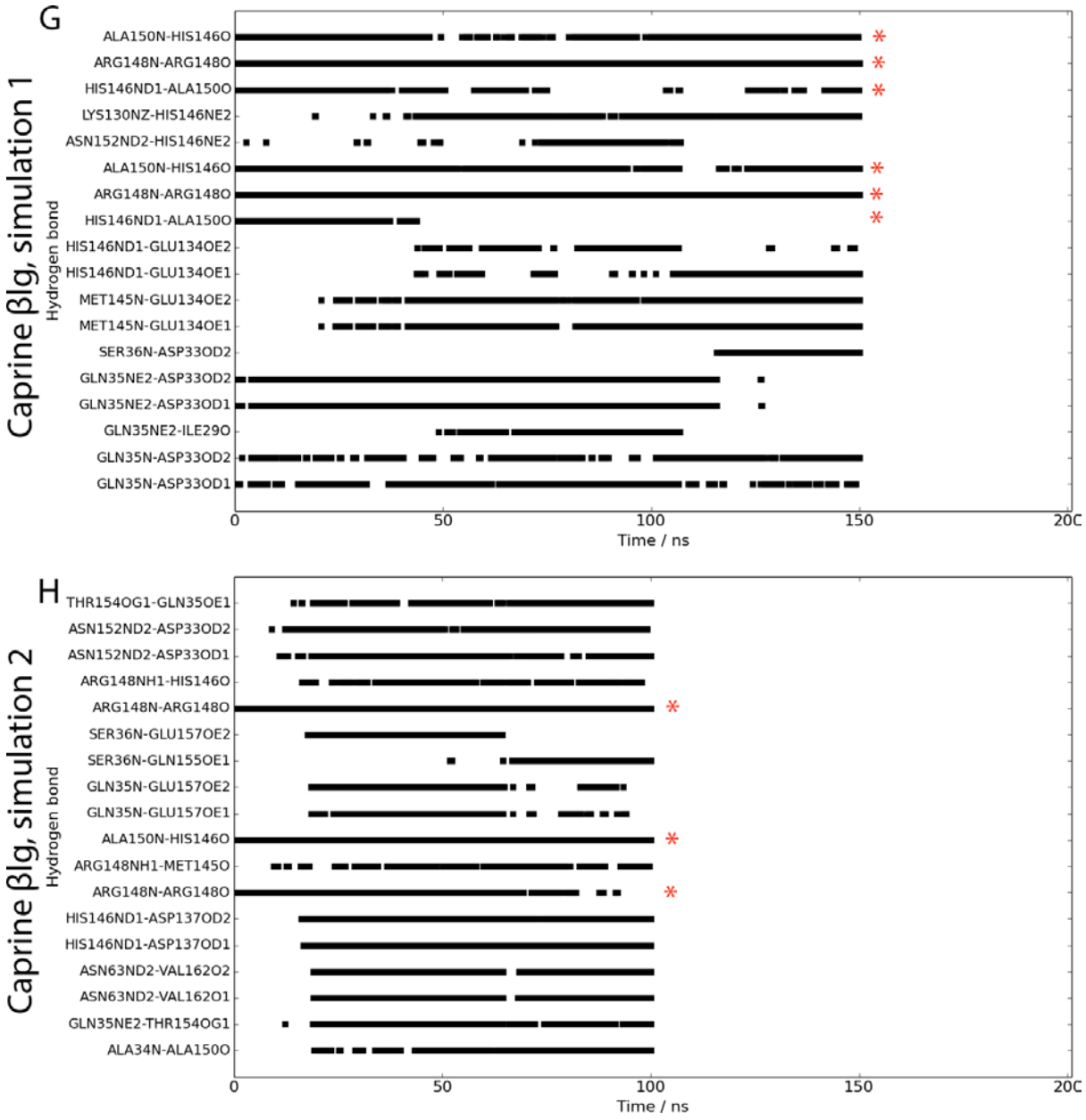


Figure 5.16: Hydrogen bonds that exist between chain A and chain B of βlg proteins for at least 20 ns during the simulation. A and B) bovine βlg A (PDB ID 1BSO), C and D) bovine βlg A (PDB ID 1BSY), E and F) bovine βlg B (PDB ID 1BSQ), G and H) caprine βlg (PDB ID 4TLJ). Red asterisks indicate hydrogen bonds seen in crystal structures of βlg between the I β strands (His146 – Ser/Ala150, Arg148 – Arg148), blue asterisks indicate hydrogen bonds seen in crystal structures between the AB loops (between Asp33, Ala34 and Arg40).

of the AB loops (Figure 5.16, see the blue asterisks). There are, however, many additional hydrogen bonds that do occur and these are all (with the exception of Thr18-Glu45, seen when

bovine βlg A dissociates) in the vicinity of these elements. In fact they nearly all involve residues within the range of 28 – 40 (near the AB loop) or 145 – 155 (near the I β-strand).

In simulations at 37°C in solution, it appears that the sidechains of Asp33 and Arg40 of opposite subunits are not always close enough to form salt-bridges (< 0.4 nm) (Figure 5.17). However, new salt bridges form during the course of the simulation as subunit re-arrangement occurs (Figure 5.18). These have not been seen in crystal structures of bovine or caprine βlg and likely compensate for the loss of interactions between Asp33 and Arg40 residues. Two of these new interactions involve residues located on the I β-strand and AB loop (Arg148 and Asp28, respectively). The rest involve the sidechains of Asp, Glu, Lys and Arg residues located on the α-helices or the GH loops (Glu127, Asp130 (Lys130 in caprine βlg), Glu131, Glu134, Lys135, Asp137, Lys138, Lys141 and Glu112, Glu114, respectively). These elements all lie near the crystallographic dimer interface (Figure 5.19). This suggests that there is considerable flexibility in the subunit arrangement and, while a similar dimer interface is observed in different crystal structures of βlg, an ensemble of possible conformations exist in solution. As observed by Bello *et al*. (2016) a combination of crystallographic data and molecular dynamics simulations allow us to observe all of the interactions that contribute to dimer stability.

In the structures that undergo complete dimer dissociation (bovine βlg A simulations 1 and 3), new salt bridges and hydrogen bonds form involving groups at sites distant from the crystallographic interface. Only one new hydrogen bond (Thr18-Glu45) can be seen for the first simulation of bovine βlg A. This is due to the length of time that a hydrogen bond is required to be present (20 ns) for the interaction to be included in the analysis. The third simulation of bovine βlg A also involves dimer dissociation, however this occurs at the very end of the simulation thus no new interfaces were present for more than 20 ns. Charged groups only need to come within 0.4 nm of one another to be included in the salt-bridge analysis. Salt bridges involving residues distant from the crystallographic dimer interface are shown for the third simulation of bovine βlg A in Figure 5.18 C (Glu51-Lys138, Lys70-Asp130 and Lys70-Glu134). However, they are not shown for bovine βlg A simulation 1 due to the sheer number of interactions that occur (95 interactions). This number is likely due to the fact that the subunits tumble past each other following dissociation and while many charged groups come within 0.4 nm of one another, the interactions between them are not long-lasting.

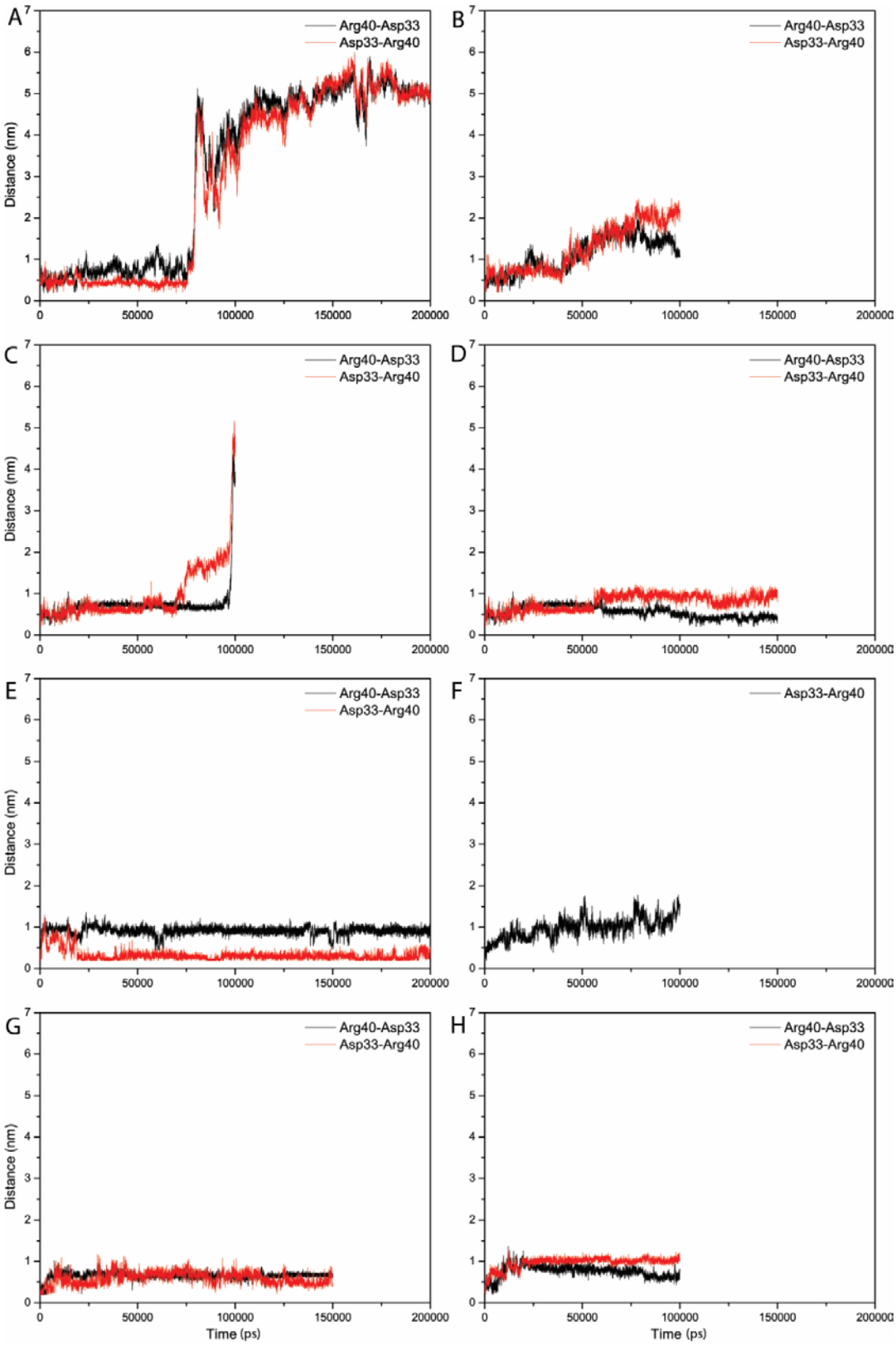


Figure 5.17: The distances between Asp33 and Arg40 residues with time. A and B) bovine βlg A (PDB ID 1BSO), C and D) bovine βlg A (PDB ID 1BSQ), E and F) bovine βlg B (PDB ID 1BSY), G and H) caprine βlg (PDB ID 4TLJ). A distance cut-off of 0.4 nm was chosen, such that groups that are never within that distance during the course of the simulation (such as for Arg40 – Asp33 in F) are not shown.

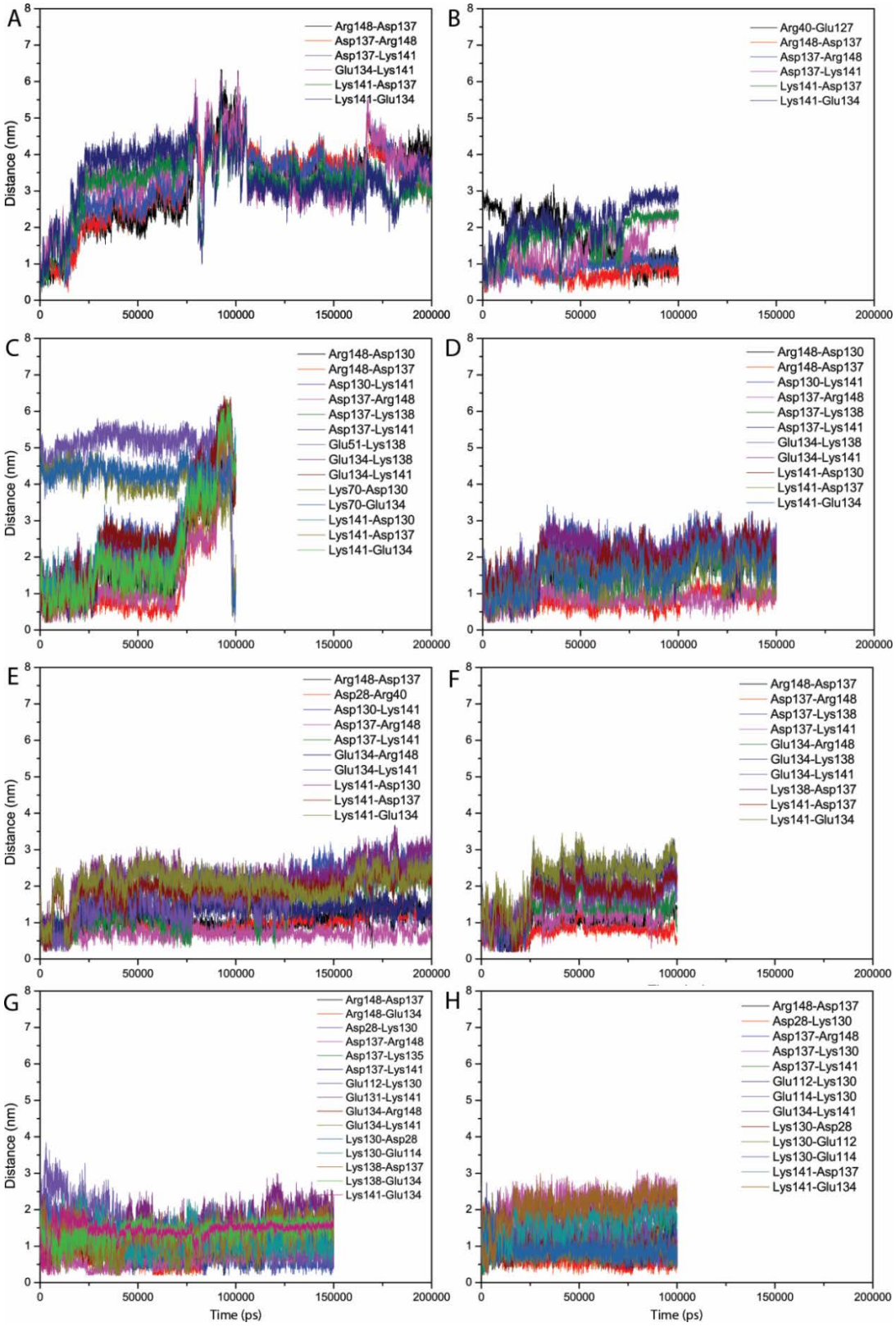


Figure 5.18: The distances between charged groups during simulations of βlg. Only groups from opposite subunits that come within 0.4 nm during the simulation are shown. A and B) bovine βlg A (PDB ID 1BSO), C and D) bovine βlg A (PDB ID 1BSQ), E and F) bovine βlg B (PDB ID 1BSY), G and H) caprine βlg (PDB ID 4TLJ).

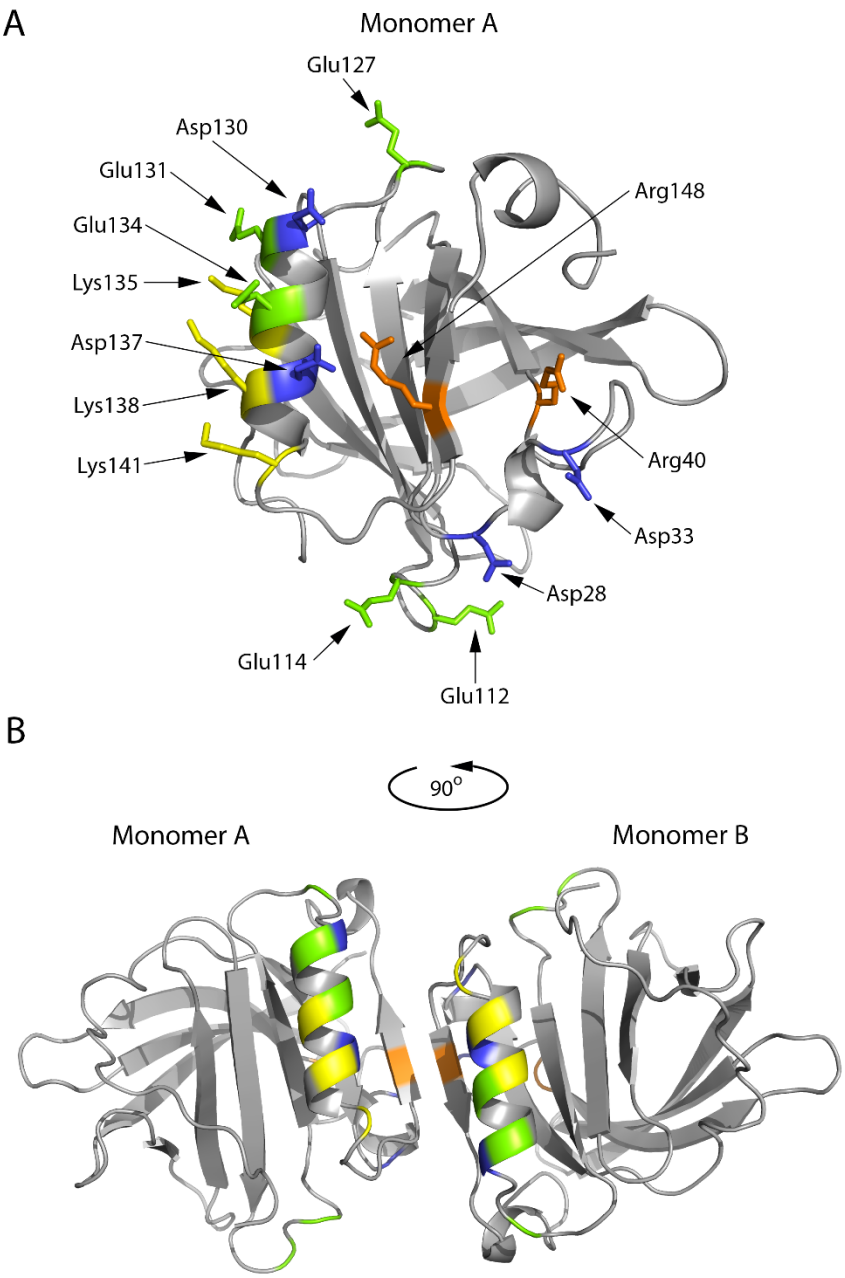


Figure 5.19: Residues that participate in salt bridge interactions during the course of the simulations. A) A single subunit is shown with participating residues shown as sticks. B) A dimer is shown to indicate the vicinity of these residues to the dimer interface.

5.3 Summary

It has been shown here that caprine βlg exhibits a similar level of resistance towards urea denaturation as the B variant of bovine βlg. These share a slightly reduced level of resistance toward urea denaturation than the A variant of bovine βlg. This suggests that the amino acid

substitutions between bovine βlg B and caprine βlg have little effect on the interactive forces holding these proteins together. Caprine βlg contains the same arrangement of disulfide bonds as the bovine variants (see section 4.5), thus it follows that it also exhibits resistance toward thermal denaturation at low pH values. This is significant as these proteins may be exposed to low pH environments during digestion.

The main difference that is seen between the stability of orthologues of βlg is dependent on the pH at which thermal denaturation is conducted. This can be rationalised by the fact that caprine βlg contains a lesser negative charge than the bovine A and B variants of βlg, and thus the isoelectric point of caprine βlg occurs at higher pH values. Heat treatment of caprine βlg resulted in visible protein aggregation at pH values below 6.9, while no precipitation was visible for bovine samples down to pH 6.5. This is likely due to disulfide-exchange induced aggregation followed by hydrophobically-driven associations between the aggregates.

These findings have significance in the processing of goat milk and goat milk products. While a heat stability dependence of goat milk on pH has already been observed (Anema and Stanley, 1998; Zadow *et al*., 1983), it may be that these observations have more to do with the properties of βlg than previously thought. It appears that these non-covalent interactions first require the formation of disulfide linked aggregates as no precipitation was seen when the thiol group of βlg was reacted with CPM dye. Blocking the free thiol group of βlg while in milk may reduce the propensity of goat milk to precipitate on heating at pH values below 6.9.

The simulated dynamics of bovine and caprine βlg do not suggest that caprine βlg is significantly more susceptible to digestive enzymes than the bovine variants. The enhanced digestibility of caprine βlg may simply be a consequence of the enhanced digestion of goat milk, due to differences in the fat and protein composition. The simulations do, however, provide us the insight that the dimer interface observed in crystal structures is a transient conformation and one of several possible in solution. Whether the dissociation seen for bovine βlg A is significant will be difficult to conclude until a greater number of simulations have been carried out for other orthologues.

Together these results indicate that the changes in primary sequence between bovine and caprine βlg can affect the physicochemical properties of these proteins. However, rather than

affecting the conformational stability or dynamic nature, they do this by altering the resultant net charge on the protein. We have seen here that this can affect the response of these proteins to heat treatment at certain pH values, while in chapter 4 it was seen to alter the association behaviour of caprine βlg at low pH. It is possible that these net charge changes may also alter the interactions of caprine βlg with other milk components, digestive enzymes or even antibodies.

5.4 References

Alexandrov, A., Mileni, M., Chien, E., Hanson, M., Stevens, R., 2008. Microscale Fluorescent Thermal Stability Assay for Membrane Proteins. Structure 16, 351–359.

Almaas, H., Cases, A., Devold, T.G., Holm, H., Langsrud, T., Aabakken, L., Aadnoey, T., Vegarud, G.E., 2006a. *In vitro* digestion of bovine and caprine milk by human gastric and duodenal enzymes. International Dairy Journal 16, 961–968.

Almaas, H., Holm, H., Langsrud, T., Flengsrud, R., Vegarud, G.E., 2006b. *In vitro* studies of the digestion of caprine whey proteins by human gastric and duodenal juice and the effects on selected microorganisms. The British journal of Nutrition 96, 562–9.

Anema, S., Stanley, D., 1998. Heat-induced, pH-dependent behaviour of protein in caprine milk. International dairy journal 8, 917–923.

Bello, M., Fragoso-Vázquez, M.J., Correa Basurto, J., 2016. Energetic and conformational features linked to the monomeric and dimeric states of bovine BLG. International Journal of Biological Macromolecules 92, 625–636.

Bello, M., Gutiérrez, G., García-Hernández, E., 2012. Structure and dynamics of β-lactoglobulin in complex with dodecyl sulfate and laurate: A molecular dynamics study. Biophysical Chemistry 165, 79–86.

Bouhallab, S., Leconte, N., Graet, L.Y., Garem, A., 2002. Heat-induced coagulation of goat milk: modification of the environment of the casein micelles by membrane processes. Le Lait 82, 673–681.

Brownlow, S., Cabral, J., Cooper, R., Flower, D., Yewdall, S., Polikarpov, I., North, A., Sawyer, L., 1997. Bovine β-lactoglobulin at 1.8 Å resolution — still an enigmatic lipocalin. Structure 5, 481–495.

Ceballos, L., Morales, E., Martínez, L., Extremera, F., Sampelayo, M., 2009. Utilization of nitrogen and energy from diets containing protein and fat derived from either goat milk or cow milk. Journal of Dairy Research 76, 497–504.

Creamer, L.K., Plowman, J.E., Liddell, M.J., Smith, M.H., Hill, J.P., 1998. Micelle stability: kappa-casein structure and function. Journal of Dairy Science 81, 3004–12.

Crowther, J., Lassé, M., Suzuki, H., Kessans, S., Loo, T., Norris, G., Hodgkinson, A., Jameson, G., Dobson, R., 2014. Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin. FEBS Letters 588, 3816–3822.

Dong, A., Matsuura, J., Allison, S.D., Chrisman, E., Manning, M.C., Carpenter, J.F., 1996. Infrared and circular dichroism spectroscopic characterization of structural differences between beta-lactoglobulin A and B. Biochemistry 35, 1450–7.

Evoli, S., Guzzi, R., Rizzuti, B., 2014. Molecular simulations of β-lactoglobulin complexed with fatty acids reveal the structural basis of ligand affinity to internal and possible external binding sites. Proteins Struct Funct Bioinform 82, 2609–2619.

Galani, D., Apenten, R., 1999. Heat-induced denaturation and aggregation of β-Lactoglobulin: kinetics of formation of hydrophobic and disulphide-linked aggregates. International Journal of Food Science & Technology 34, 467–476.

Gutiérrez-Magdaleno, G., Bello, M., Portillo-Téllez, C.M., Rodríguez-Romero, A., García-Hernández, E., 2013. Ligand binding and self-association cooperativity of β-lactoglobulin. Journal of Molecular Recognition 26, 67–75.

Kella, N.K., Kinsella, J.E., 1988. Enhanced thermodynamic stability of beta-lactoglobulin at low pH. A possible mechanism. The Biochemical journal 255, 113–8.

Kuwata, K., Era, S., Hoshino, M., Forge, V., Goto, Y., Batt, C., 1999. Solution structure and dynamics of bovine β-lactoglobulin A. Protein Science 8, 2541–2545.

Manderson, G., Creamer, L., Hardman, M., 1999. Effect of heat treatment on the circular dichroism spectra of bovine beta-lactoglobulin A, B, and C. Journal of Agricultural and Food Chemistry 47, 4557–67.

Manderson, G., Hardman, M., Creamer, L., 1998. Effect of Heat Treatment on the Conformation and Aggregation of β-Lactoglobulin A, B, and C. Journal of Agricultural and Food Chemistry 46, 5052–5061.

Montilla, A., Calvo, M., 1997. Goat's milk stability during heat treatment: effect of pH and phosphates. Journal of Agricultural and Food Chemistry 45, 931–934.

Oliveira, K.M., Valente-Mesquita, V.L., Botelho, M.M., Sawyer, L., Ferreira, S.T., Polikarpov, I., 2001. Crystal structures of bovine β-lactoglobulin in the orthorhombic space group C2221. European Journal of Biochemistry 268, 477–484.

Qin, B., Bewley, M., Creamer, L., Baker, H., Baker, E., Jameson, G., 1998a. Structural Basis of the Tanford Transition of Bovine β-Lactoglobulin. Biochemistry 37, 14014–14023.

Qin, B., Creamer, L., Baker, E., Jameson, G., 1998b. 12-Bromododecanoic acid binds inside the calyx of bovine β-lactoglobulin. FEBS Letters 438, 272–278.

Qin, B., Jameson, G., Bewley, M., Baker, E., Creamer, L., 1999. Functional implications of structural differences between variants A and B of bovine β-lactoglobulin. Protein Science 8, 75–83.

Reddy, I.M., Kella, N.K., Kinsella, J.E., 1988. Structural and Conformational Basis of the Resistance of B-Lactoglobulin to Peptic and Chymotryptic Digestion. Journal of Agricultural Food Chemistry 36, 737–741.

Rivas, G., Fernández, J., Minton, A., 2001. Direct observation of the enhancement of noncooperative protein self-assembly by macromolecular crowding: Indefinite linear self-association of bacterial cell division protein FtsZ. Proceedings of the National Academy of Sciences 98, 3150–3155.

Roefs, S., Kruif, K., 1994. A Model for the Denaturation and Aggregation of β-Lactoglobulin. European Journal of Biochemistry 226, 883–889.

Uhrínová, S., Smith, M.H., Jameson, G.B., Uhrín, D., Sawyer, L., Barlow, P.N., 2000. Structural changes accompanying pH-induced dissociation of the beta-lactoglobulin dimer. Biochemistry 39, 3565–74.

Wang, C., Zhu, Y., Wang, J., 2015. Comparative study on the heat stability of goat milk and cow milk. Indian Journal of Animal Research 50, 610–613.

Xu, D., Tsai, C., Nussinov, R., 1997. Hydrogen bonds and salt bridges across protein-protein interfaces. Protein Engineering, Design and Selection 10, 999–1012.

Yi, C., Wambo, T.O., 2015. Factors affecting the interactions between beta-lactoglobulin and fatty acids as revealed in molecular dynamics simulations. Physical Chemistry Chemical Physics 17, 23074–80.

Zadow, J.G., Hardham, J.F., Kocak, H.R., Mayes, J.J., 1983. The stability of goat's milk to UHT processing. The Australian Journal of Dairy Technology 1, 20–23.

Chapter Six

6. Interactions between β-lactoglobulin and milk components

6.1 Introduction

6.1.1 A definitive role for βlg is yet to be ascribed

Despite being intensively studied for over half a century, a convincing physiological role for βlg is yet to be proposed. As a member of the lipocalin family, a group of proteins with a wide variety of ligand-binding related functions (see section 1.7.1), it is natural to assume that βlg also has a role involving the binding and transport of a particular ligand. Early on it was demonstrated that bovine βlg can bind vitamin A (Pérez and Calvo, 1995) . This led to the hypothesis that βlg functions as a transporter of vitamin A between mother and offspring. However, since then a vast multitude of small hydrophobic molecules have been shown to be capable of being bound by bovine βlg (see section 1.7.2). The apparent lack of selectivity makes it less likely that βlg is a specific fatty acid or vitamin transporter, although, it is still possible that βlg may act in the protection of these sorts of ligands from oxidation or by increasing the bioavailability of these ligands (le Maux *et al*., 2014; Pérez *et al*., 1992) .

The observation of ligand binding *in vitro*, using proteins in isolation, does not necessarily translate to a physiological role *in vivo*. There is a need to observe the interactions of βlg in its natural surroundings, i.e. within milk. This is not a straightforward task due to the complexity of milk as a solution. Fortunately, the recent development of fluorescence-detection analytical

ultracentrifugation has made it possible to explore the interactions of biomolecules within complex solutions (Cole *et al.*, 2008) .

6.1.2 Fluorescence-detection analytical ultracentrifugation

Analytical ultracentrifugation (AUC) is an effective method used to characterise the solution behaviour of biomolecules, such as proteins. The concentration distribution of proteins in solution during centrifugation can be monitored over time using either the absorbance or Rayleigh interference optical systems (Kroe and Laue, 2009) . However, to observe the sedimentation of a particular biomolecule within complex solutions requires a unique feature to be inherent in that molecule, such that it may be selectively observed. It may be that the molecule of interest absorbs at a certain wavelength where the other components of the system do not absorb. If not naturally occurring, this can be engineered into the molecule of choice by adding a tag. The sedimentation of the molecule on its own can be compared to the behaviour seen in the presence of other molecules. If an interaction occurs, this will affect the sedimentation of that molecule.

If one has an analytical ultracentrifuge equipped with a fluorescence detection system, a fluorescent tag may be added to the biomolecule of choice. The detection system utilises confocal optics to track the location of fluorophores during sedimentation (Kroe and Laue, 2009) . The system exhibits high sensitivity and high selectivity (Kroe and Laue, 2009) , which allows the observation of minor components within complex solutions, for example serum, cell lysate and, of importance here, milk.

6.1.3 Overview

The aim of this chapter is to characterise the interactions (if any) that occur between caprine and bovine βlg within their physiological environments: cow and goat milk. In order to do this, a fluorescent label (fluorescein isothiocyanate, FITC) was attached to each of these proteins. This allowed the sedimentation of these βlg proteins to be observed within the complex milieu of milk utilising fluorescence-detection AUC. This is the first time that the interactions of βlg within milk have been examined using this technique.

When compared to the sedimentation behaviour of just βlg in a simple buffer, it can be seen that both bovine and caprine βlg behave very differently within milk. It appears that these proteins form a complex with another, much larger, component. These interactions are not the result of the self-association of βlg into a higher molecular mass species, or due to the addition of the FITC label. Due to the size of the complex, it was hypothesised that βlg may be interacting with an immunoglobulin protein. The role of βlg may be to bind and protect immunoglobulins so that they may traverse the stomach undigested and thus become available for absorption in the intestine. This may confer protective immunity to the infant ruminant, and provides a new explanation for the abundance of βlg in milk.

6.2 Results and discussion

6.2.1 Part 1: Bovine and caprine βlg interact with a component in cow and goat milk

6.2.1.1 Fluorescent labelling of βlg proteins

In order to analyse the sedimentation behaviour of bovine βlg A and caprine βlg in milk, a fluorescent tag (FITC) was attached to each protein. Thus, the sedimentation of these proteins can be selectively observed against a complex background of biomolecules.

Bovine βlg A and caprine βlg were incubated in the presence of FITC in order to obtain fluorescently-labelled proteins (as in section 2.6.3.2). FITC can bind to any of the available amine groups, such as the N-terminus or side-chains of lysine residues, on the surface of the protein. There are 15-16 lysine residues per monomer of bovine and caprine βlg, respectively. A labelling ratio of 0.3 – 1 mole of FITC per 1 mole of protein was deemed an acceptable level of labelling. The self-association behaviour of both bovine and caprine βlg is unaffected by the addition of FITC. At the low concentrations (0.15 – 0.65 μM) examined by AUC using the fluorescence detection system, both FITC-labelled bovine βlg and caprine βlg are predominantly monomeric (Figure 6.1 A and B). When compared at similar protein concentrations (10 – 30 μM, Figure 6.1 C and D) labelled and unlabelled βlg proteins behave similarly (i.e. they are in a monomer-dimer self-association with K_D values within the range of 5 – 45 μM, see section 4.2.1).

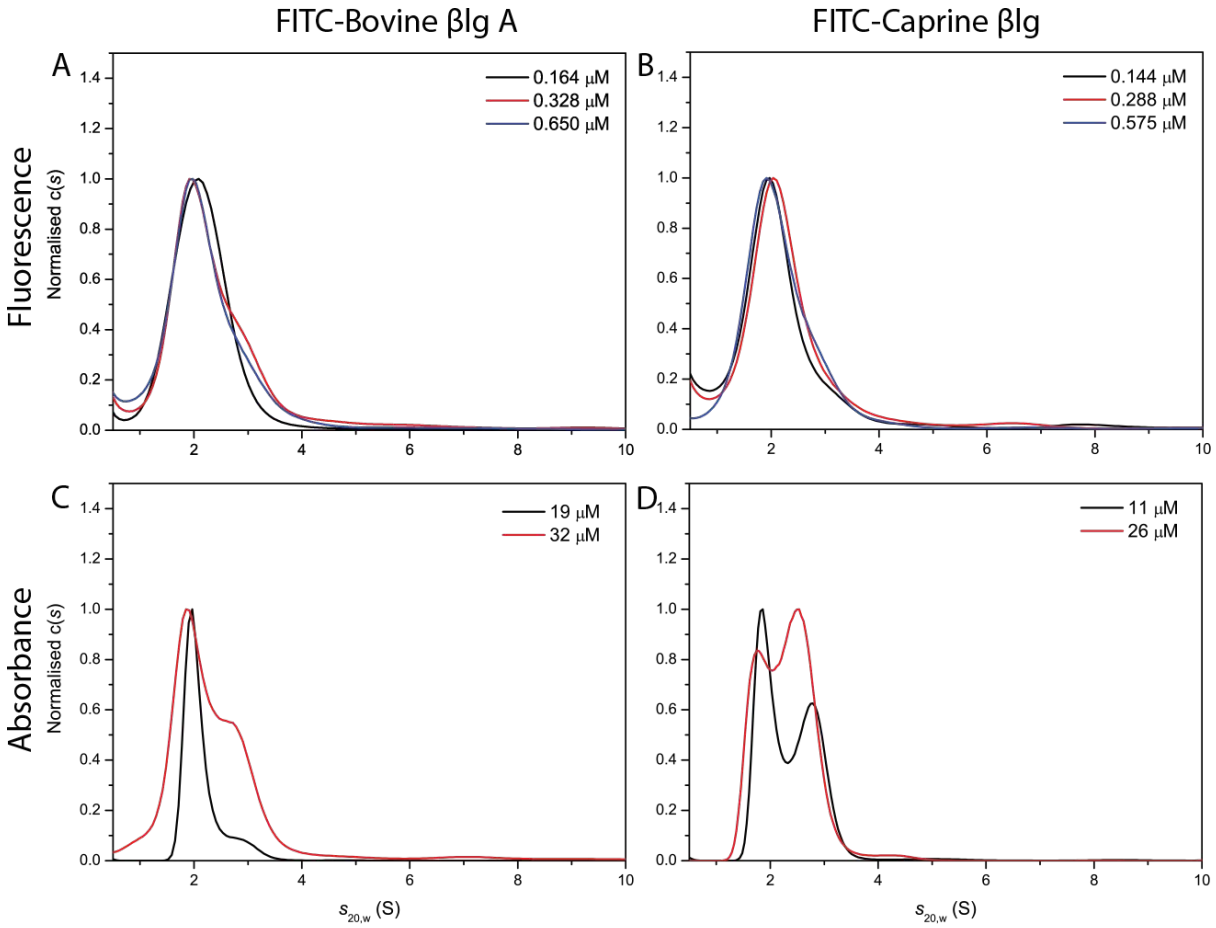


Figure 6.1: Sedimentation velocity analysis of FITC-labelled bovine βlg A and caprine βlg. A and B utilised the fluorescence-detection optical system while C and D utilised the absorbance optical system.

6.2.1.2 The sedimentation of βlg in milk

The sedimentation of fluorescently-labelled bovine βlg A and caprine βlg is altered significantly in the presence of cow or goat milk. This suggests that an interaction forms between βlg and another component within milk. A dilution series of cow and goat milk samples were prepared by diluting with buffer (0.01 M Tris, 0.1 M NaCl, pH 7) and fluorescently-labelled βlg proteins were added to each sample to a final concentration of 0.75 μM. This is much lower than the concentration of βlg that is usually found in milk (~150 μM (le Maux *et al.*, 2014)). The sedimentation of FITC-labelled βlg was observed by monitoring the distribution of fluorescence at 488 nm within the AUC cell during centrifugation. When no milk is present (i.e. buffer and protein only) bovine and caprine βlg are predominantly monomeric with a small amount of

dimer present (see the violet line in Figure 6.2). Once milk is added to the solution, the sedimentation profile changes considerably, with the appearance of a new species at a much higher s value of ~8 S (Figure 6.2). The continuous sedimentation distribution ($c(s)$) plots in Figure 6.2 have been normalised to the height of the peak at 2 S, i.e. βlg that is free in solution. This emphasises that as the concentration of milk increases the proportion of βlg that is found in complex (i.e. the peak at 8 S) increases.

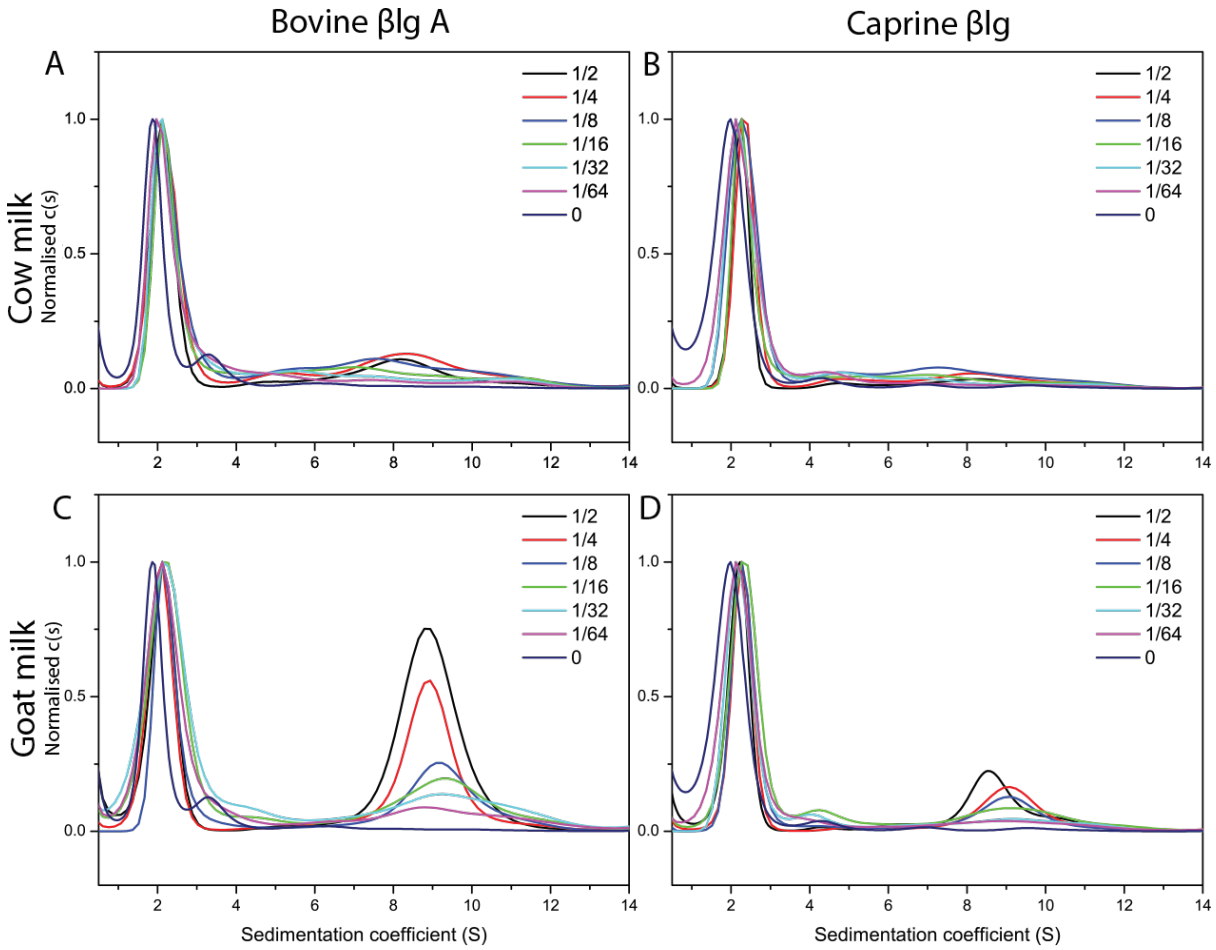


Figure 6.2: Sedimentation velocity analysis of bovine and caprine βlg in cow and goat milk. Each milk was serially diluted from ½ to 1/64 dilution. Protein was added to a final concentration of 0.75 µM. A) FITC-labelled bovine βlg A in cow milk, B) FITC-labelled caprine βlg in cow milk, C) FITC-labelled bovine βlg A in goat milk, D) FITC-labelled caprine βlg in goat milk.

6.2.1.3 FITC does not interact with anything in milk

The sedimentation of FITC itself is not altered in the presence of milk (Figure 6.3), which rules out the possibility that the FITC dye molecule is interacting with a component in milk. The reactivity of the dye molecule was first quenched by reacting with Tris. In both buffer and milk, minimal sedimentation is observed (as can be seen in the sedimentation data shown in Figure 6.3 B), owing to the small size of FITC (389 Da). This suggests that FITC on its own is not capable of forming a complex with any components in milk, and is clearly not responsible for the peak at 8 S in Figure 6.2

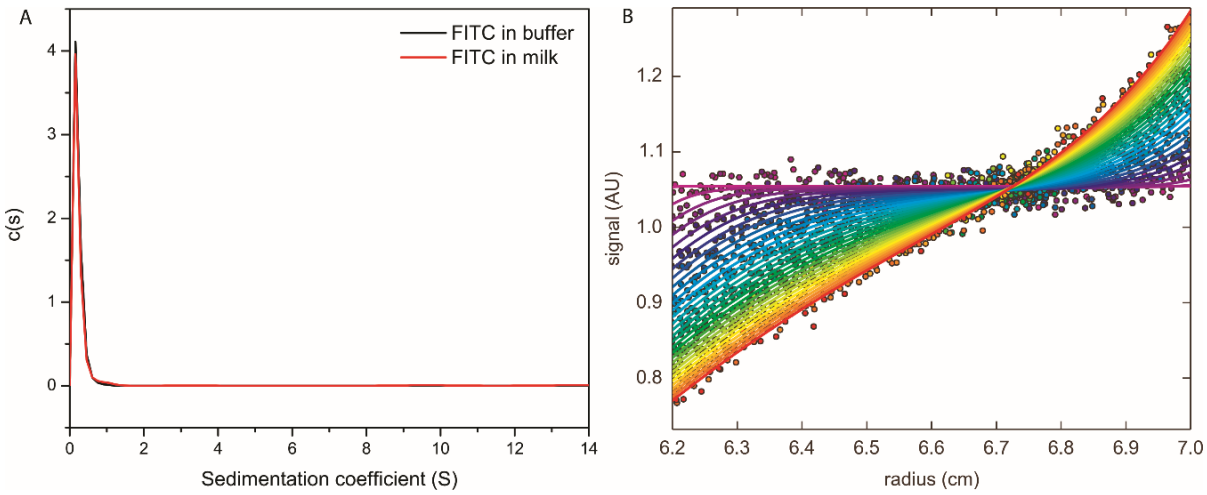


Figure 6.3: A) Sedimentation velocity analysis of FITC (15 µM) in buffer (0.01 M Tris, 0.1 M NaCl, pH 7) and in goat milk (1/20 dilution). B) The sedimentation observed for FITC in goat milk. Note that this is characteristic of the limited sedimentation seen for very small particles.

6.2.1.4 Self-association behaviour of βlg proteins at high concentrations

It is possible that the 8 S species is a product of self-association of βlg in the complex solution of milk where the concentration of βlg is high (150 µM (le Maux *et al*., 2014)). It is possible that the fluorescently-labelled βlg is capable of mixing with the endogenous βlg already present in the milk samples, which would effectively increase the concentration of fluorescently-labelled βlg.

Sedimentation velocity experiments of bovine βlg A and caprine βlg at high concentrations were performed to rule out the possibility of self-association of βlg into a higher molecular mass species. These unequivocally show a single species with an s value of around 2.6 over the concentration range examined (100, 200 and 400 µM), consistent with a dimer (Figure 6.4). This proves that βlg does not form any higher-order species above a dimer, even at protein concentrations that may be encountered in milk (~150 µM (le Maux *et al*., 2014)).

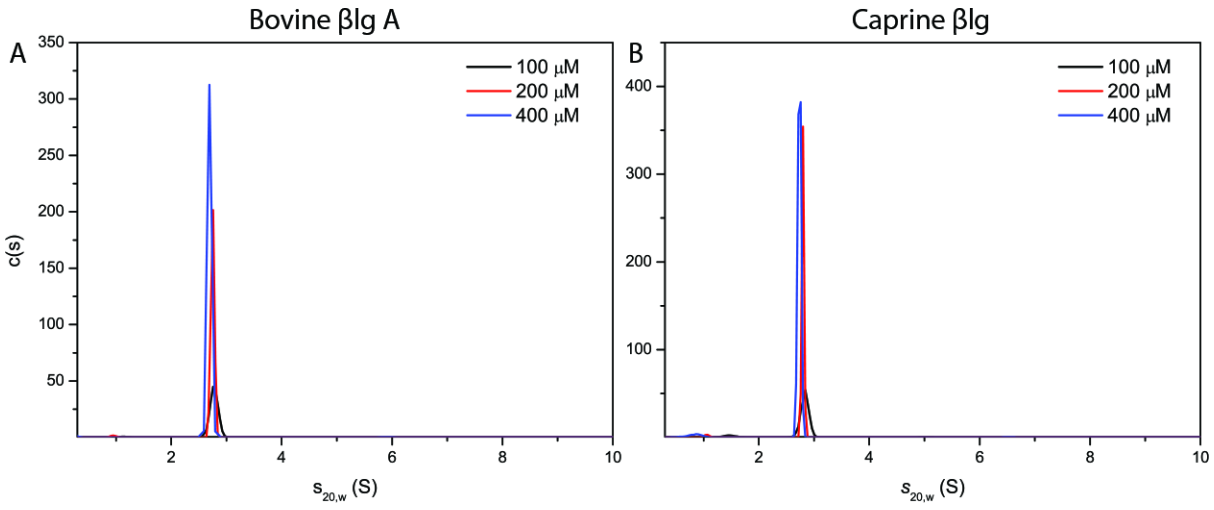


Figure 6.4: Sedimentation velocity analysis utilising interference optics to look at high concentrations (up to 400 µM) of A) bovine βlg A and B) caprine βlg. The analysis of βlg at high concentrations requires the use of the Rayleigh interference optical system due to the limits of the absorbance optical system.

The fact that the weight-averaged s of FITC-labelled βlg increases with increasing milk concentration is indicative of an interaction forming between βlg and a component in milk, in a concentration-dependent manner (Figure 6.5). The weight-averaged s values were obtained by integrating the *c*(*s*) distributions, seen in Figure 6.2, between s values of 1 and 13. The levelling off of the weight-averaged s values seen at fractions of milk of 0.125 and above may be due to the viscosity and density of the samples increasing as the milk concentration increases. Unfortunately, it was not possible to measure the viscosity and density of each sample in this case and the values were thus held constant for each during data analysis. An

increase in the viscosity and density of samples would make the same sized particle travel slower and would appear as a smaller s value.

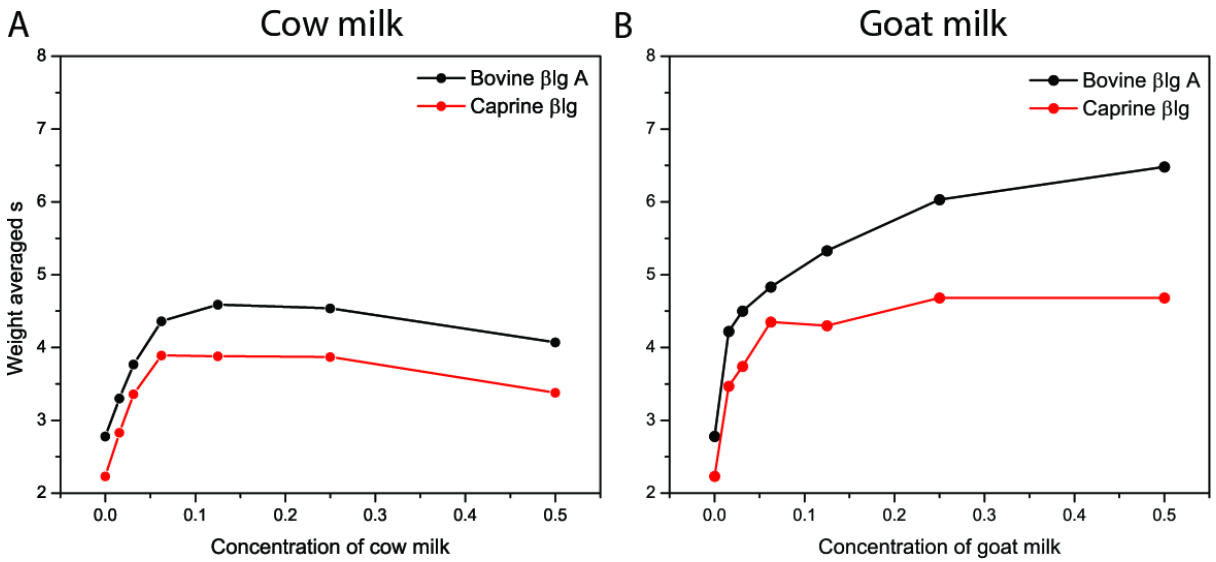


Figure 6.5: The weight-averaged s of FITC-labelled bovine βlg A and caprine βlg as a function of cow milk (A) and goat milk (B) concentration.

In a similar analysis, in order to keep the viscosity and density the same between samples, a consistent dilution of milk (¼) was used across all samples and instead the protein concentration was altered (from 0.156 to 5 µM). Interestingly, as the concentration of βlg increases, the proportion of βlg that is seen in complex decreases (Figure 6.6). This can be seen most clearly for Figure 6.6 C, by observing the decrease in the peak height at 8 S as the protein concentration increases from 0.156 µM to 5 µM.

One explanation for the decrease in the proportion of βlg found in complex with increasing βlg concentrations, is that the interacting component within milk is the factor that is limiting complex formation. That is, at this concentration of milk (¼), even at the lowest βlg concentration (0.156 µM), most of the binding sites within the milk are occupied. As more βlg is added to the solution it contributes only to the pool of free βlg. This would also explain the decrease in the weight-averaged s with βlg concentration (Figure 6.6 E and F).

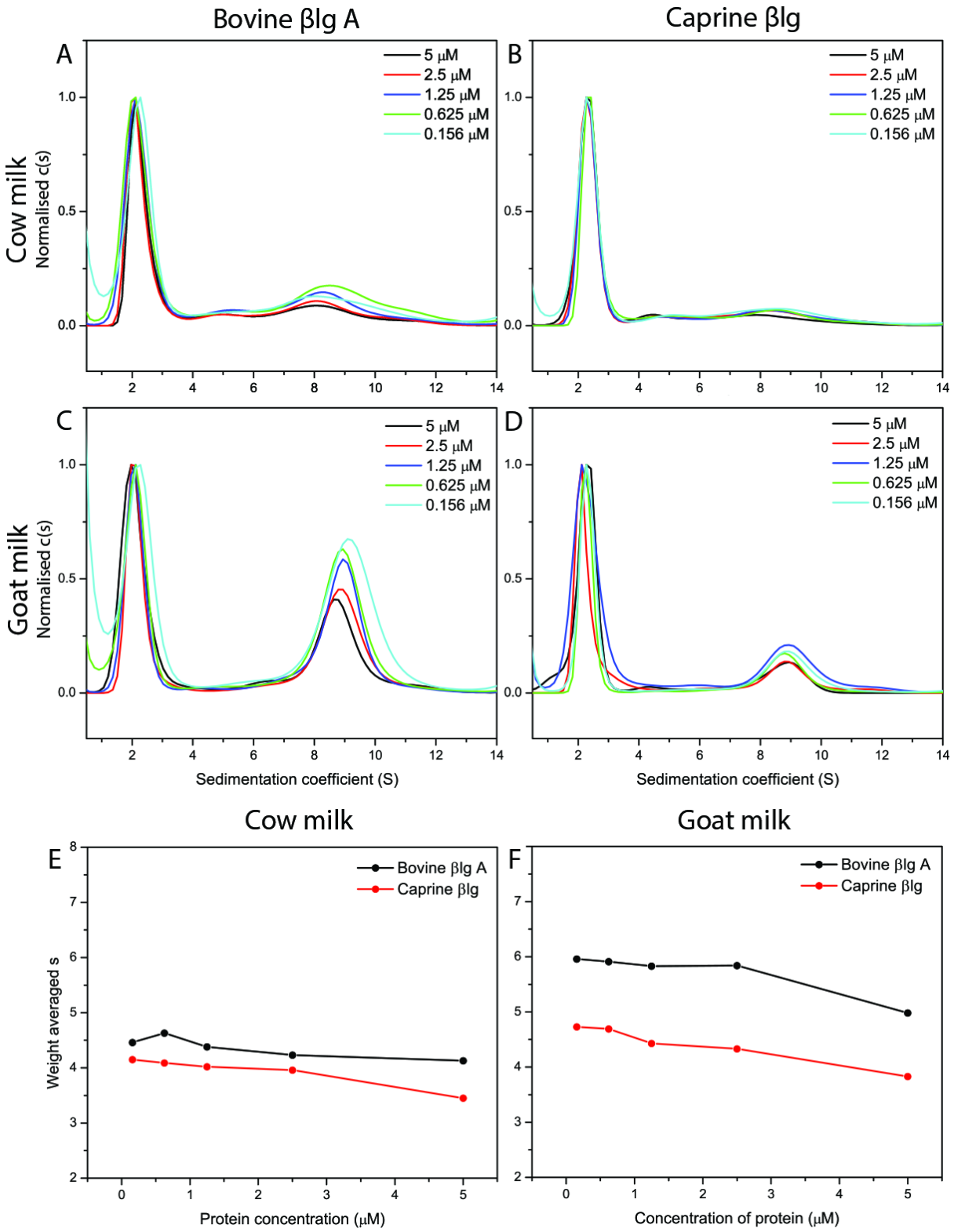


Figure 6.6: Sedimentation velocity analysis of bovine and caprine βlg in cow and goat milk. The concentration of milk for each for each sample was a ¼ dilution of skimmed milk. The concentration of protein ranged from 0.156 to 5 μM. A) FITC-labelled bovine βlg A in cow milk, B) FITC-labelled caprine βlg in cow milk, C) FITC-labelled bovine βlg A in goat milk, D) FITC-labelled caprine βlg in goat milk, E and F) weight-averaged s as a function of βlg concentration.

Interestingly, we see different behaviours between both the type of milk used and the orthologue of βlg examined. In goat milk we see a greater amount of complex formation as compared to that seen in cow milk (i.e. the 8 S peaks are higher in samples of goat milk than in cow milk, compare Figure 6.2 C and D to A and B, and Figure 6.6 C and D to A and B). This may suggest that the interacting component is present in goat milk at a higher concentration, and therefore has a higher number of available binding sites. We also see a greater amount of complex forming for bovine βlg A than for caprine βlg (compare Figure 6.2 A and C to B and D, and Figure 6.6 A and C to B and D). As the same concertation of βlg was added to each sample, this may mean that bovine βlg A has a higher binding affinity for the interacting component in milk than caprine βlg, which allows a greater proportion of this protein to bind and form a complex.

6.2.1.5 Unlabelled βlg is capable of competing for binding with labelled βlg

A competition assay was performed to assess whether unlabelled βlg could complete for binding with FITC-labelled βlg for complex formation. If so, this would decrease the proportion of FITC-labelled βlg that is seen in complex. This would mean that the binding interaction is specifically between βlg and the component within milk and not between FITC and the component, or due to any structural changes induced in βlg by the addition of FITC. Unlabelled βlg was added to samples of milk containing a fixed amount of FITC-labelled βlg. As the concentration of unlabelled βlg is increased, the amount of FITC-labelled βlg involved in the complex decreases, shown by a reduction in peak height at 8 S (Figure 6.7 A and B) and a decrease in the weight-averaged s of FITC-labelled βlg with increasing unlabelled βlg (Figure 6.7 C).

Curiously, the position of the higher s peak shifts towards the left with increasing concentrations of unlabelled βlg (seen most clearly for caprine βlg in Figure 6.7 B), suggesting that the complex is decreasing in size. It is still unclear why this may be, however, it may again be due to an increase in the viscosity and density of the solution as a high concentration of unlabelled protein is added. This increase in viscosity and density would have a greater effect on the sedimentation of the larger molecule due to this molecule encountering a greater

frictional force, which would explain why the position of the peak corresponding to free βlg remains unaffected.

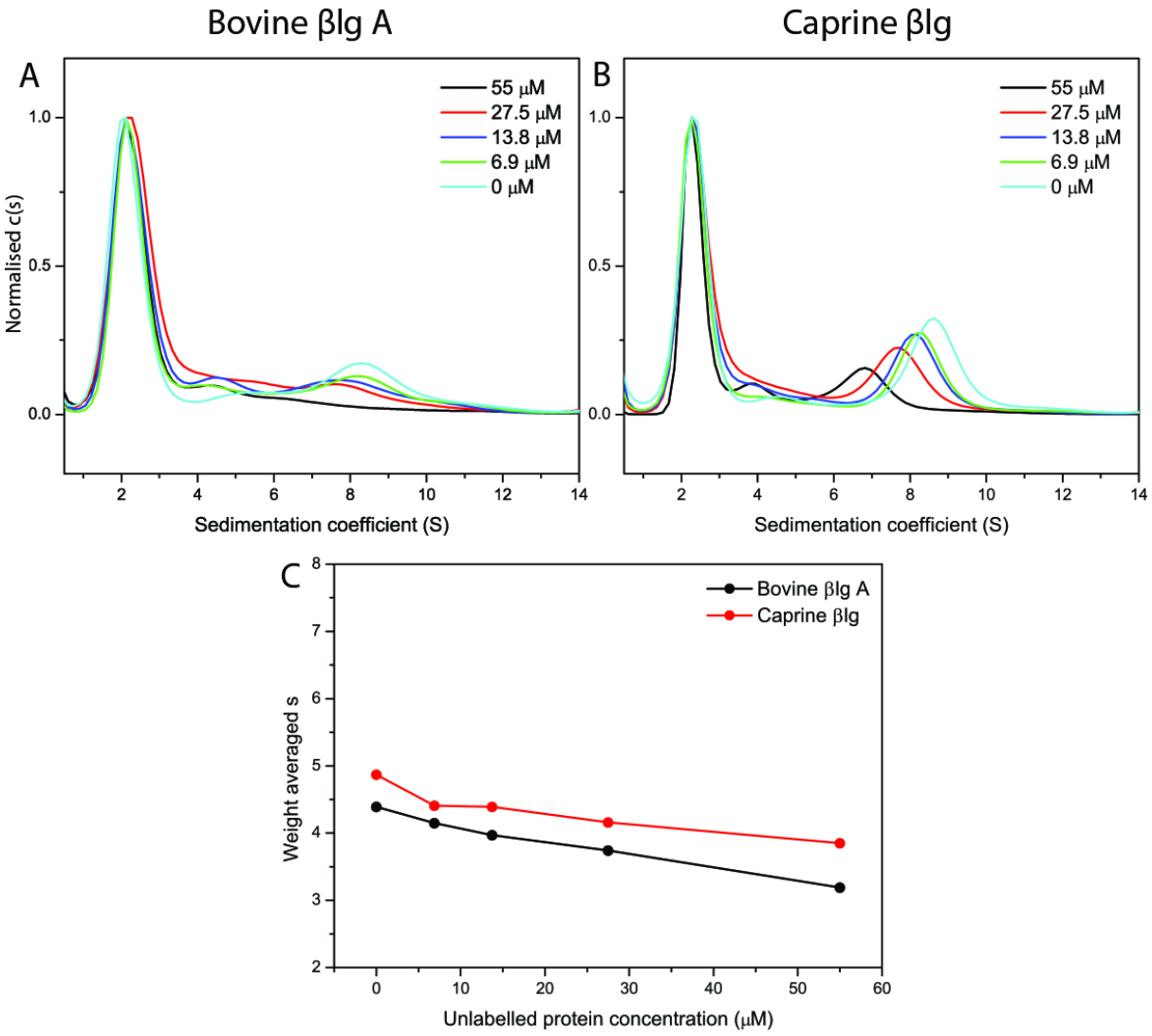


Figure 6.7: Sedimentation velocity analysis of FITC labelled βlg proteins (0.75 μM) in milk (¼ dilution) with increasing concentrations of unlabelled βlg. A) FITC-labelled bovine βlg A in cow milk with unlabelled bovine βlg A, B) FITC-labelled caprine βlg in goat milk with unlabelled caprine βlg, C) weight-averaged s as a function of unlabelled βlg concentration.

6.2.1.6 βLg is present in milk in two populations

Note: The experimental work contained in section 6.2.1.6 was carried out by researchers at AgResearch Ltd, Ruakura under the guidance of Dr Alison Hodgkinson, co-supervisor for this project.

In order to see if the higher molecular mass species of βlg can be observed in milk, a sample of skimmed cow milk was analysed by gel-filtration chromatography. This technique separates molecules based on size, with large molecules eluting first and smaller molecules eluting later. βLg was identified in the eluted fractions by means of a Western dot-blot utilising antibodies specific for bovine βlg. It is apparent that βlg elutes in two places: once at the location expected due to its molecular mass (see Box A in Figure 6.8), and once more at a location expected for a much higher molecular mass species (Box B, Figure 6.8). This suggests that βlg is present in milk as two populations of different sized species, which agrees directly with what is seen in AUC experiments. The importance of this result lies in the fact that this experiment involves βlg that is endogenous to the milk that has been sampled; it has not been recombinantly expressed or modified in any way, yet we see the same outcome.

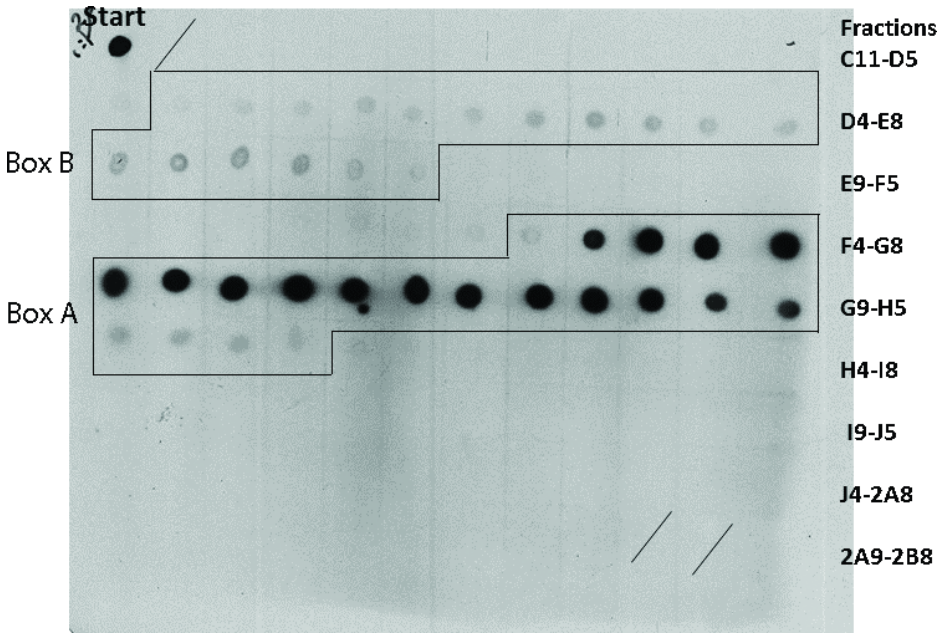


Figure 6.8: Western dot-blot of cow milk fractions following size-exclusion utilising anti-bovine βlg antibodies. Bovine βlg elutes according to its molecular mass in fractions F4-G8, G9-H5 and H4-I8, and also in a region much higher than its expected molecular mass: D4-E8, E9-F5.

In summary, complex formation has been observed between bovine and caprine βlg and an unknown component within cow and goat milk, utilising both analytical ultracentrifugation and size exclusion chromatography. The possibility of self-association of βlg into a higher-molecular weight species is unlikely due to the lack of any species larger than a dimer at concentrations higher than those observed in milk, and due to the lack of any well-defined intermediate species observed between s values of 2 and 8 S. It is also unlikely that the FITC dye molecules attached to the surface of βlg proteins are responsible for the complex formation seen. This has been shown by the fact that the sedimentation of FITC is unaltered in the presence of milk. Even more convincingly, we still see the formation of a higher molecular mass species of βlg, even when native βlg is examined in its physiological environment.

6.2.2 Part 2: Identification of the binding partner within milk

While it has been shown that bovine and caprine βlg can both form associations with κ-casein in milk, this interaction is heat-induced and involves disulfide-exchange (Creamer *et al*., 1998; Henry *et al*., 2002) . Thus, the 8 S species seen here in milk in AUC and size-exclusion experiments is unlikely to be an association between βlg and κ-casein. In an effort to identify components of milk that would be large enough to form the 8 S species, it was hypothesised that it may be an immunoglobulin protein that βlg is interacting with. Monomeric immunoglobulins are around 160 kDa in size (Hurley and Theil, 2011) .

The immunoglobulins present in milk represent a history of the antigens that that individual has been exposed to, and that their immune system has responded to, during their lifetime (Hurley and Theil, 2011) . Immunoglobulins are secreted into milk via receptor-mediated mechanisms and transferred to offspring during suckling. Immunoglobulins then enter the digestive tract of the infant where, if they are able to resist digestion, they are available for uptake by the infant or to provide protection within the gastrointestinal tract (Hurley and Theil, 2011) .

Immunoglobulins can be separated into several classes including IgM, IgA, IgG, IgE and IgD (Hurley and Theil, 2011) . IgG, IgA and IgM are the major immunoglobulins found in colostrum and milk. IgM appears initially when an organism is first exposed to an antigen. IgA is the major

immunoglobulin found in mucosal secretions towards preventing mucosal infections, while IgG is the predominant immunoglobulin found in bovine milk (Hurley and Theil, 2011) . Monomers of immunoglobulins contain an antigen-binding fragment (Fab), which includes a variable amino acid domain, and a constant fragment (Fc), which has a constant amino acid sequence among molecules of the same sub-class (Hurley and Theil, 2011) . While IgG are monomeric, IgA can form dimers and IgM form pentamers (Hurley and Theil, 2011) .

Interestingly, in humans and rabbits, two species that both lack βlg in their milk, IgG is transferred mainly to the foetus via the placenta (Hurley and Theil, 2011) . Conversely, the offspring of ungulates (such as cows and goats) are born agammaglobulinemic (i.e. with no circulating antibodies) and thus rely on the uptake of immunoglobulins, especially IgG, from colostrum and milk for systemic immune protection (Hurley and Theil, 2011) . It follows that the IgG content in human colostrum is lower than that in other species and is similar to that found in human milk (Hurley and Theil, 2011) . The predominant immunoglobulin in human colostrum is IgA which provides protection in the gastrointestinal tract of the infant (Cakebread *et al*., 2015) . In bovine and caprine colostrum the levels of IgG are very high (greater than 75% of the total immunoglobulin content, compared to 2% in human colostrum) and decrease with successive milkings (Crosson *et al*., 2010) . This IgG provides systemic immune protection to the calf, while the lower levels of IgA provide protection in the gastrointestinal tract (Hurley and Theil, 2011) .

IgA is transported into mucosal secretions via the polymeric immunoglobulin receptor (pIgR). IgA is then released, but retains a portion of the pIgR, termed the secretory component, forming secretory IgA (sIgA). The secretory component protects sIgA against proteolytic degradation within the gastrointestinal tract (Cakebread *et al*., 2015) . Thus, IgG is less resistant to degradation than sIgA (Crottet and Corthésy, 1998) . Ungulate species rely solely on the uptake of IgG from colostrum and milk to strengthen their immune system. It is hypothesised here that bovine and caprine βlg are capable of interacting with IgG within cow and goat milk. The role of βlg may therefore be to increase the resistance of IgG toward digestion by proteolytic enzymes and/or denaturation at the lowered pH conditions encountered in the gastrointestinal tract during digestion.

6.2.2.1 Immunoglobulins co-elute from milk with βlg

Note: The experimental work contained in section 6.2.2.1 was carried out by researchers at AgResearch Ltd, Ruakura under the guidance of Dr Alison Hodgkinson, co-supervisor for this project.

When the fractions of milk eluted from size exclusion chromatography in section 6.2.1.6 were probed with anti-bovine IgG, IgA and IgM antibodies it became apparent that these immunoglobulins elute from milk at the same time as the higher molecular mass bovine βlg (Figure 6.9). This evidence is consistent with the hypothesis that bovine βlg binds immunoglobulins in milk to form the higher molecular-mass complex seen in AUC experiments.

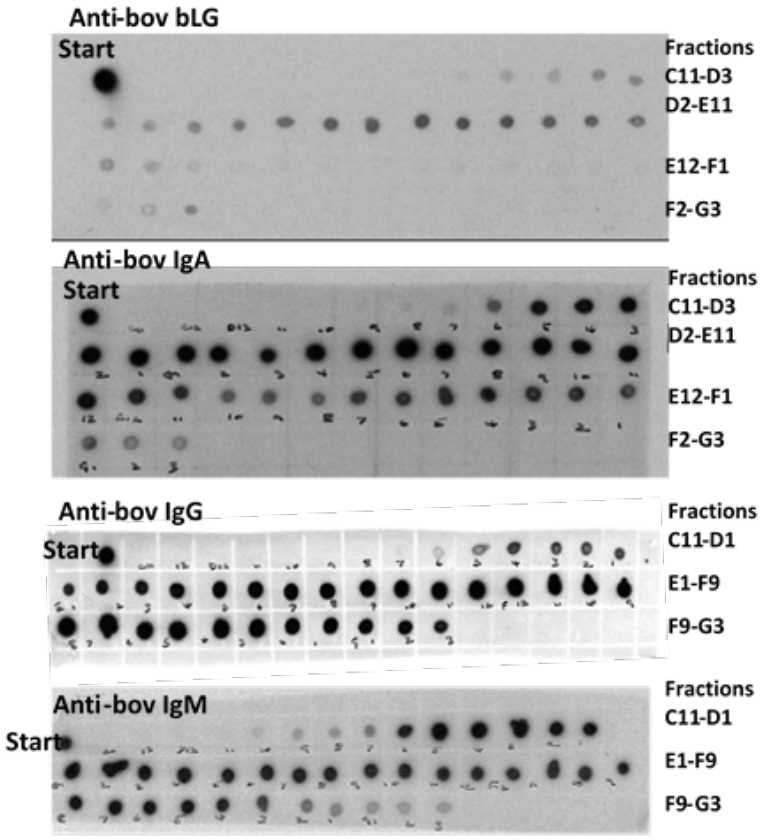


Figure 6.9: Western dot-blot of milk fractions following size-exclusion chromatography utilising anti-bovine βlg, anti-bovine IgA, anti-bovine IgG and anti-bovine IgM.

6.2.2.2 Bovine, but not caprine, βlg can be seen to interact with IgG in isolation

Bovine and caprine IgG proteins, purified from the serum of non-immunised animals, were purchased in order to characterise the interaction that may be occurring between these proteins and βlg. The sedimentation of bovine βlg appears to be influenced by the addition of bovine IgG, while no interaction is observed between caprine βlg and caprine IgG.

For the following analyses the use of the absorbance optics were necessitated as I no longer had access to a fluorescence-detection system AUC. The sedimentation of caprine βlg in the presence of purified caprine IgG was analysed by observing the absorbance at 280 nm along the length of the AUC cell. However, both βlg and IgG will contribute to the absorbance at 280 nm, thus it is difficult to tell whether the sedimentation of caprine βlg is altered in the presence of IgG or if the resulting *c*(*s*) distribution is simply a combination of the signals from the individual proteins (Figure 6.10).

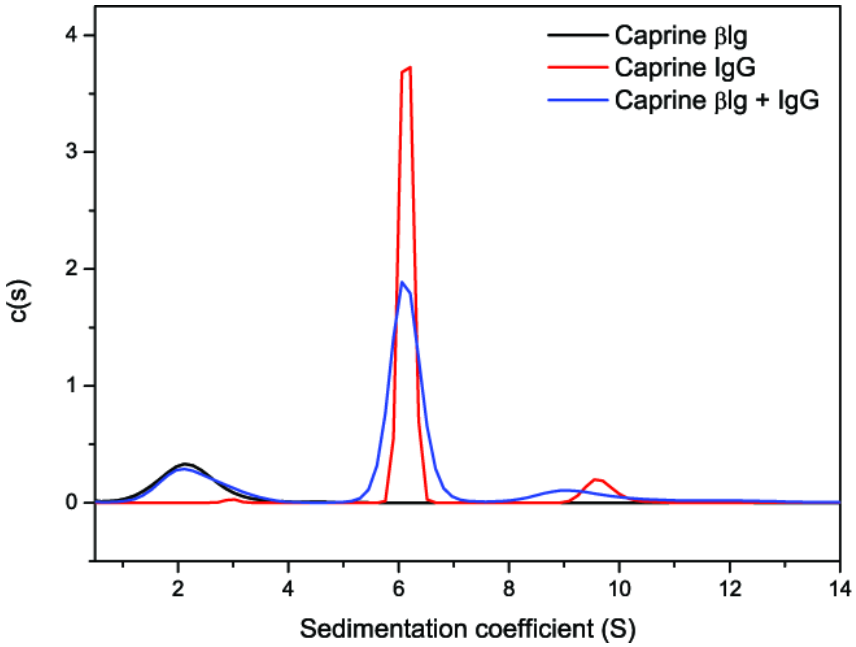


Figure 6.10: Sedimentation velocity analysis of caprine IgG (5 µM) and caprine βlg (5 µM) in buffer (50 mM NaCl, 50 mM sodium phosphate, pH 7) and in combination.

While the sensitivity is not as comparable to the fluorescence detection system, the absorbance at 495 nm using the AUC absorbance optical system can also be used to selectively monitor the sedimentation of a FITC-labelled protein within a mixture of proteins. Initial experiments were performed in order to ascertain that complex formation could be visualised for FITC-labelled βlg in milk using the absorbance optics. A 1/25 dilution of goat milk allowed a complex to be observed for FITC-labelled bovine and caprine βlg at a FITC concentration of 8 µM (protein concentrations of 19 and 11 µm) (Figure 6.11). The weight-averaged s values of bovine and caprine βlg increased from 2.08 to 3.83 and 2.27 to 3.55 in buffer and in milk, respectively. This means that the absorbance optical system can successfully be used to further investigate the interactions that involve βlg proteins in milk.

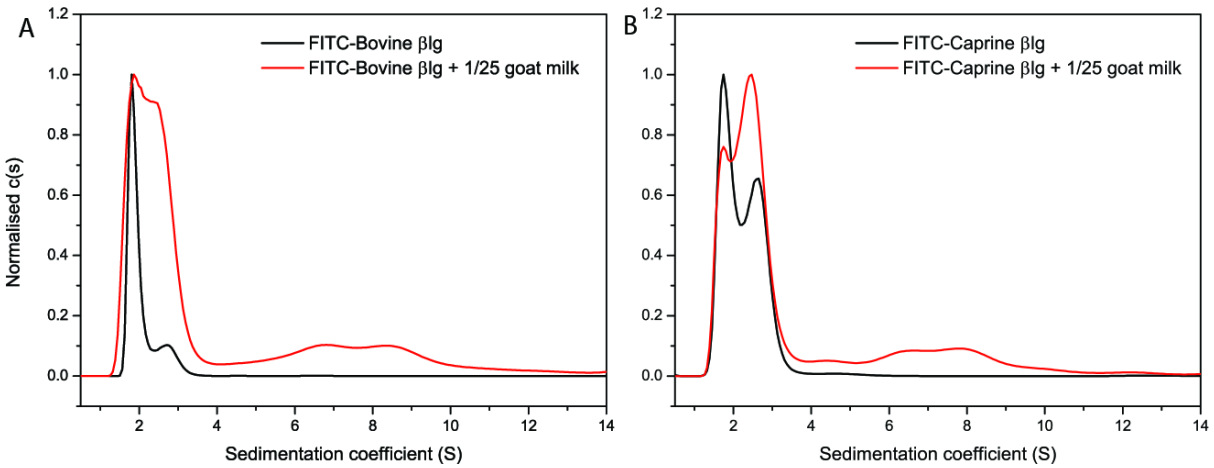


Figure 6.11: Sedimentation velocity analysis of A) FITC-labelled bovine βlg in buffer and in milk, and B) FITC-labelled caprine βlg in buffer and in milk.

The sedimentation of bovine βlg A and caprine βlg was analysed in the presence of purified bovine and caprine IgG. The sedimentation of bovine βlg A is affected by the addition of IgG (Figure 6.12 A, note the change in sedimentation behaviour between the black line (without IgG) and the red line (with IgG)), suggesting that an interaction occurs between these two proteins. However, no interaction between caprine βlg and caprine IgG is observed (Figure 6.12 B). The weight averaged s of FITC-labelled bovine βlg A increases from 2.12 to 2.59 in the presence of bovine IgG, while the weight-averaged s of FITC-labelled caprine βlg only increases

slightly with increasing concentrations of caprine IgG (2.40 to 2.51). The concentration of both bovine βlg and bovine IgG (both 30 µM) was higher than for the caprine counterparts (12.5 and 3.75 – 15 µM, respectively), which may explain why a greater interaction was seen between the bovine proteins.

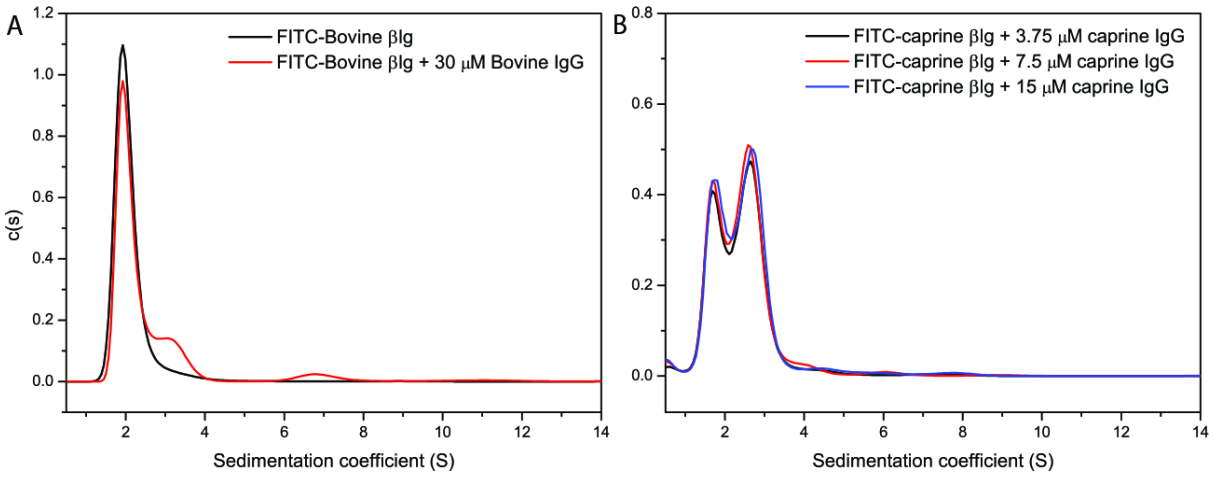


Figure 6.12: Sedimentation velocity analyses of βlg proteins in the presence of IgG. A) Bovine βlg (30 µM) and bovine IgG (30 µM), B) Caprine βlg (12.5 µM) and caprine IgG (3.75 – 15 µM).

It was thought that the interaction may be encouraged by adding milk to the sample, as there may be other components within milk required for complex formation. The weight-averaged s of FITC-labelled bovine βlg A increases from 2.12 in buffer (see Figure 6.12 A) to 2.93 in the presence of goat milk (1/200 dilution, see Figure 6.13 A). Adding IgG to this sample increases the weight-averaged s further from 2.93 to 3.43 (Figure 6.13). This suggests that the addition of milk increases the amount of complex formed, but this may be due to the increase in concentration of βlg and IgG already present in the milk. On the other hand, no effect was seen for the addition of caprine IgG to FITC-labelled caprine βlg in goat milk (final milk dilution of 1/25) (weight-averaged s = 3.78, 3.80, 3.77 and 3.68 for 0, 3.75, 7.5 and 15 µM IgG, respectively).

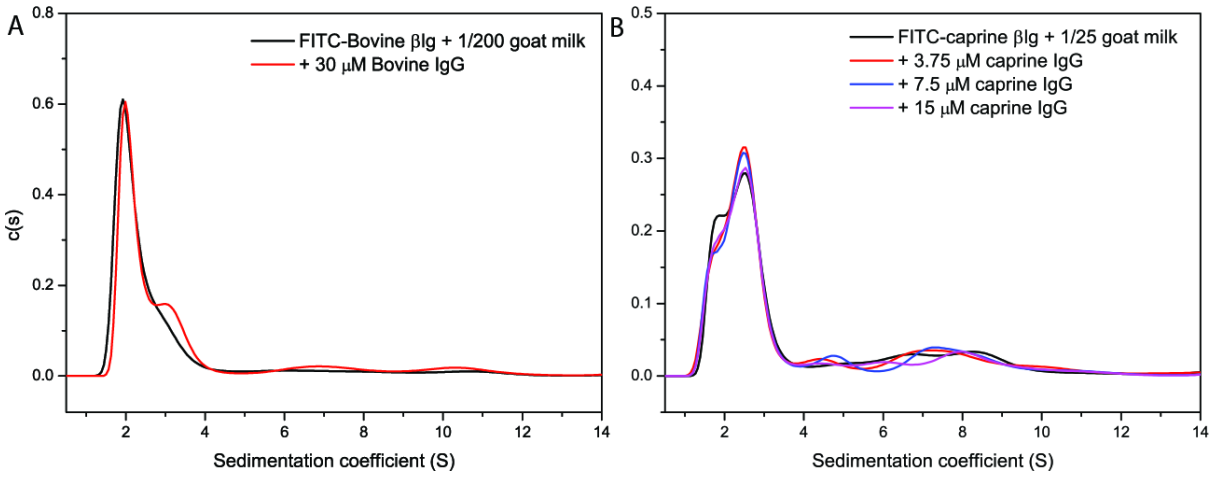


Figure 6.13: Sedimentation velocity analyses of βlg proteins in the presence of goat milk and IgG. A) Bovine βlg (30 μM) and bovine IgG (30 μM), B) Caprine βlg (12.5 μM) and caprine IgG (3.75 – 15 μM).

No evidence of interaction is apparent when FITC-labelled bovine IgG was analysed in the presence of unlabelled bovine or caprine βlg (Figure 6.14). The weight-averaged s value of FITC-labelled bovine IgG varies from 6.77 in buffer to 6.69, 6.73 and 6.65 in the presence of increasing concentrations of bovine βlg, and to 6.77, 6.77 and 6.70 in the presence of increasing concentrations of caprine βlg. In this case the interaction may be difficult to discern given that the attachment of βlg (~36 kDa) will not produce a large size differential from non-interacting FITC-labelled bovine IgG (~160 kDa on its own (Hurley and Theil, 2011)). Investigating higher concentrations while utilising fluorescently labelled βlg proteins is likely a better approach going forward to characterise the interaction between βlg and IgG proteins.

The two main sub-classes of IgG in serum are IgG1 and IgG2 (Hurley and Theil, 2011). While these are present in an approximately 1:1 ratio in serum, IgG1 is found in much higher concentrations in bovine colostrum and milk (Hurley and Theil, 2011). It may be that βlg is involved in specific interactions with only particular subclasses of IgG. This may explain why we only see a limited amount of interaction between βlg and IgG purified from serum.

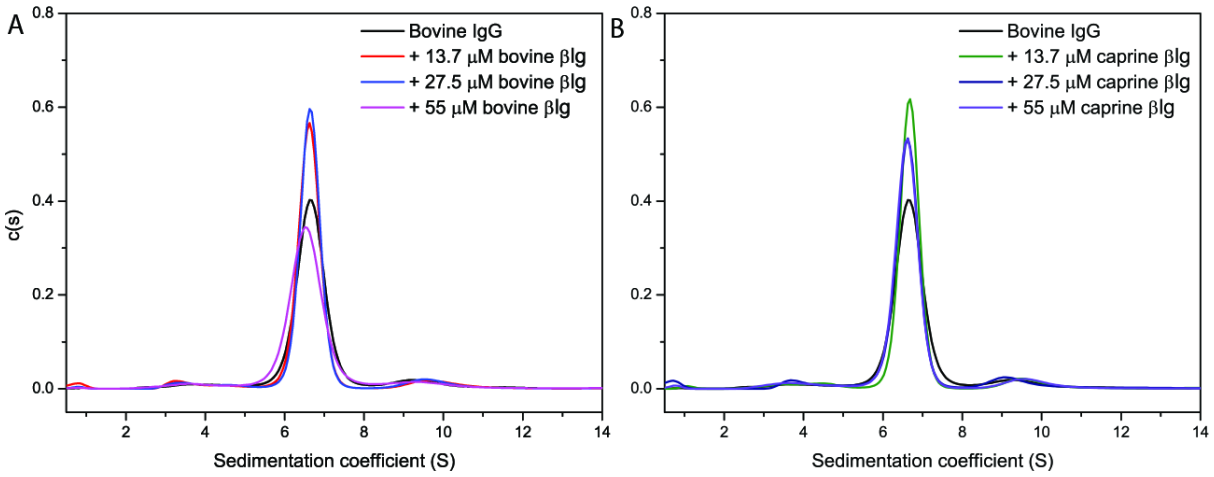


Figure 6.14: Sedimentation velocity analysis of FITC-labelled bovine IgG in the presence of A) bovine βlg and B) caprine βlg.

6.3 Summary

Here I have presented for the first time, to my knowledge, evidence of a protein-protein interaction that occurs in milk between βlg and another milk component. This work has demonstrated the utility of fluorescence-detection analytical ultracentrifugation for observing biomolecular interactions within complex solutions. The sedimentation of both bovine and caprine βlg are significantly altered in the presence of milk. I have shown that this interaction is not due to the interaction of a milk component with FITC, nor is it due to FITC-induced structural changes incurred in the βlg structure.

Due to the approximate size of the βlg complex formed it was hypothesised that the interacting component may be an immunoglobulin. The co-elution of βlg with immunoglobulins IgG, IgA and IgM from cow milk following size-exclusion chromatography supports my hypothesis. Analytical ultracentrifugation analyses indicate that the sedimentation of bovine βlg is altered in the presence of bovine IgG. This is further evidence of an interaction forming between βlg and an immunoglobulin.

Given the abundance of βlg in ruminant milk, and its absence in the milk of humans, it is unlikely that this protein would fulfil a role which is still necessary in humans. While human neonates obtain IgG via the placenta during pregnancy, ungulates (such as cows and goats) rely on the

absorption of these immunoglobulins from milk. Thus, it is perceivable that the function of βlg may be related to the transport of IgG between mother and offspring. βLg may function to protect these immunoglobulins, particularly IgG, from digestive enzymes as they traverse the digestive tract, leaving them available for absorption by the ruminant offspring.

6.4 References

Cakebread, J., Humphrey, R., Hodgkinson, A., 2015. Immunoglobulin A in Bovine Milk: A Potential Functional Food? Journal of Agricultural and Food Chemistry 63, 7311–7316.

Cole, J.L., Lary, J.W., P Moody, T., Laue, T.M., 2008. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. Methods in cell biology 84, 143–79.

Creamer, L.K., Plowman, J.E., Liddell, M.J., Smith, M.H., Hill, J.P., 1998. Micelle stability: kappa-casein structure and function. Journal of Dairy Science 81, 3004–12.

Crosson, C., Thomas, D., Rossi, C., 2010. Quantification of Immunoglobulin G in Bovine and Caprine Milk Using a Surface Plasmon Resonance-Based Immunosensor. Journal of Agricultural and Food Chemistry 58, 3259–3264.

Crottet, P., Corthésy, B., 1998. Secretory component delays the conversion of secretory IgA into antigen-binding competent F(ab')2: a possible implication for mucosal defense. Journal of immunology 161, 5445–53.

Henry, G., Mollé, D., Morgan, F., Fauquant, J., Bouhallab, S., 2002. Heat-induced covalent complex between casein micelles and beta-lactoglobulin from goat's milk: identification of an involved disulfide bond. Journal of Agricultural and Food Chemistry 50, 185–91.

Hurley, W.L., Theil, P.K., 2011. Perspectives on immunoglobulins in colostrum and milk. Nutrients 3, 442–74.

Kroe, R.R., Laue, T.M., 2009. NUTS and BOLTS: Applications of fluorescence-detected sedimentation. Analytical Biochemistry 390, 1–13.

Le Maux, S., Bouhallab, S., Giblin, L., Brodkorb, A., Croguennec, T., 2014. Bovine β-lactoglobulin/fatty acid complexes: binding, structural, and biological properties. Dairy Science & Technology 94, 409–426.

Pérez, MD, Calvo, M, 1995. Interaction of β-lactoglobulin with retinol and fatty acids and its role as a possible biological function for this protein: a review. Journal of Dairy Science 78, 978–988.

Pérez, M.D., Sanchez, L., Aranda, P., Ena, J.M., Oria, R., Calvo, M., 1992. Effect of beta-lactoglobulin on the activity of pregastric lipase. A possible role for this protein in ruminant milk. Biochimica et Biophysica Acta 1123, 151–5.

Chapter Seven

7. Conclusions and future directions

7.1 Overview

The properties of caprine βlg from goat milk are not nearly as well characterised as those for bovine βlg from cow milk. This knowledge is crucial for understanding the behaviour of cow and goat milk during digestion and during processing, and to understand the underlying causes of milk protein allergies. The first aim of this thesis is to characterise and compare the physicochemical and biophysical properties of caprine βlg to those of bovine βlg in order to better understand the behaviours of cow and goat milk.

Given the abundance of βlg in both cow and goat milk, and the intriguing lack of this protein in the milk of humans (Kontopidis *et al*., 2004), it is curious that the physiological role of this protein is yet to be uncovered. The second aim of this thesis is to identify interactions that occur between βlg and milk components under physiological conditions, towards understanding the physiological relevance of this abundant whey protein.

7.2 Bovine and caprine βlg contribute to milk allergies

Milk allergies are common, with a prevalence of 2 – 3% in infants (Saarinen *et al*., 2005). Goat milk is increasing in popularity as an alternative to cow milk, particularly as a first protein source in infants, due to its perceived health benefits and lesser allergenic burden (Haenlein, 2004; Jandal, 1996; Lara-Villoslada *et al*., 2004). However, while a greater amount of goat milk is

required to elicit the same reaction as cow milk, goat milk can still cause allergic reactions, particular if the patient is already sensitised to cow milk (Lara-Villoslada *et al*., 2004). It is important to understand the behaviours and properties of each of the constituents of cow and goat milk in order to understand the immunogenic nature of these milks.

It has been shown that bovine βlg contributes to the allergic response in patients with cow milk protein allergies (Kapila *et al*., 2013). Meanwhile, very little is known about the caprine orthologue of βlg. Here, the elucidation of the crystal structure of caprine βlg reveals that this orthologue shares a very similar structure with that of the closely related bovine orthologues. This explains the ability of immune reaction-causing immunoglobulin E (IgE) antibodies that have been raised against bovine βlg to cross-react with caprine βlg. Differences in the self-association behaviour of caprine βlg at low pH values, revealed here using analytical ultracentrifugation, may alter the presentation of these proteins to the immune system during digestion.

The reduced allergenic burden of goat milk is most likely due to a lower level of protein allergens passing through the digestive tract in their native state. The conformational flexibility (as observed by means of molecular dynamics simulations) and the stability towards denaturation by a chemical denaturant, urea, are seen here to be very similar between the bovine and caprine orthologues of βlg. These findings suggest that these protein orthologues would be digested by proteolytic enzymes to a similar extent. The enhanced digestibility of caprine βlg is likely due to compositional differences between cow and goat milk, and in particular the reduced α_{s1}-casein content. This results in a softer curd formation in the stomach, which is more susceptible to enzymatic digestion (Ceballos *et al*., 2009). From my studies we now have a more in depth knowledge of the properties of caprine βlg which helps us to understand the immunogenic nature of these milks.

7.3 βLg is involved in the response of milk to heat treatment

The heat treatment of milk and dairy products is required to eliminate pathogenic bacteria. However, the fouling of heating surfaces incurred during this process can be costly and reduce yields (Pesic *et al*., 2016). The aggregation of bovine βlg during heat treatment of cow milk

plays a dominant role in this process (de Jong, 1997), thus substantial effort has been dedicated to understanding the thermal denaturation behaviour of bovine βlg. Much less is known about the behaviour of caprine βlg during heat treatment.

If cow or goat milk is heat-treated below a certain pH, significant precipitation of casein and whey proteins occur (Anema and Stanley, 1998; Zadow *et al*., 1983). For goat milk this occurs at pH values below 6.9, whereas for cow milk this occurs at pH values below 6.6. Given that the natural pH of goat and cow milk are 6.6 and 6.7, respectively, goat milk is less heat-stable than cow milk at its natural pH (Anema and Stanley, 1998). It is believed that these differences are due to altered interactions between denatured βlg and κ-casein. What is unknown is whether these interactions in cow and goat milk are affected by differences in the thermal stability of bovine and caprine βlg.

The thermal denaturation stability of bovine and caprine βlg has been assessed by monitoring the changes in the circular dichroism spectra as a function of temperature. At pH values above 6.9, caprine and bovine βlg display similar thermal denaturation behaviour and the midpoints of the melting transitions are similar between proteins. Below pH 6.9, however, precipitation occurs in samples of caprine βlg and unfolding transitions cannot be observed. Precipitation was not seen for bovine βlg at any of the pH values examined (pH 6.5 – 7.4). The precipitation seen for samples of caprine βlg is likely due to the lesser net negative charge on the surface of these proteins, which allow hydrophobically-driven associations to occur between heat-induced aggregates of caprine βlg. This knowledge is key to understanding the heat denaturation processes of cow and goat milk in order to inform the optimisation of processing conditions.

7.4 The physiological role of βlg is yet to be defined

Despite the considerable amount of effort that has been devoted to understanding the behaviour and properties of βlg, particularly of the bovine variants, the physiological role for this protein has still not been revealed. It is of interest to uncover the function of this elusive protein in order to explain the abundance of this highly-conserved protein in certain species'

milks. The fact that human milk is devoid of βlg suggests that this role is no longer required in humans.

Here an original hypothesis has been proposed for the physiological role of βlg. This is based on evidence of a protein-protein interaction occurring in milk under physiological conditions, revealed for the first time using fluorescence-detection analytical ultracentrifugation. The co-elution of immunoglobulins with βlg from milk following size-exclusion chromatography supports the hypothesis that βlg is capable of interacting with immunoglobulins, and in particular immunoglobulin G (IgG).

The transfer of immunoglobulins from a mother to her offspring provides passive immunity to the infant, which strengthens their developing immune system. Human neonates receive IgG from their mother via the placenta during pregnancy (Hurley and Theil, 2011). Cows and goats, on the other hand, are born without IgG and must obtain them from their mother's milk and colostrum (Hurley and Theil, 2011). It is possible that βlg may facilitate the transfer of IgG from mother to offspring. The increased stability of βlg at low pH values lends itself to the suggestion that it may somehow augment the resistance of IgG towards degradation within the acidic and enzymatically active digestive tract.

7.5 Future directions

The enhanced digestion of caprine βlg over bovine βlg is observed in simulated *in vitro* digestion studies of cow and goat milk (Almaas *et al*., 2006). However, the similarities between the structure and conformational stability of bovine and caprine βlg, as revealed here, suggest that these proteins will be presented in the same way to digestive proteolytic enzymes. What remains to be investigated is the simulated *in vitro* digestion of these protein orthologues in isolation. This will inform whether the amino acid substitutions between the bovine and caprine orthologues of βlg are responsible for the enhanced digestion of caprine βlg, or if this is due to the compositional differences between cow and goat milk.

It is hypothesised that the precipitation of caprine βlg at pH values of 6.7 and 6.5 are the result of the lesser net negative charge that caprine βlg carries at these pH values, compared with

bovine βlg. No precipitation was seen for samples of bovine βlg at the pH values investigated (pH 6.5 to 7.4). It will be worth investigating whether similar levels of precipitation are seen for bovine βlg upon heat treatment at pH values below pH 6.5, as this will reduce the net negative charge on the protein. However, these pH values are getting further from the natural pH of cow milk and are thus of less physiological relevance.

The most apparent unresolved question is the identity of the interacting component forming a complex with βlg in milk. Efforts will continue to unambiguously identify the interacting partner and to thoroughly characterise this protein-protein interaction. Analytical ultracentrifugation analyses carried out using higher concentrations of βlg and IgG proteins may provide further proof of an interaction between these proteins. It may be useful to utilise the immunoglobulins that co-elute with βlg in the fractions obtained by size-exclusion chromatography, in case βlg is forming specific interactions with certain populations of the immunoglobulins found in milk. In order to further characterise the interaction, cross-linking experiments will be useful to inform the stoichiometry of the proteins participating in the interaction. Ultimately, a crystal structure of the complex would provide invaluable insight into how these complexes form. If βlg is indeed involved in the transport of IgG between mother and offspring, while the production of hypoallergenic cow or goat milk that is devoid of βlg would be of benefit to humans, the implications of this for the offspring of the cows and goats involved would need to be carefully considered.

7.6 References

Almaas, H., Holm, H., Langsrud, T., Flengsrud, R., Vegarud, G.E., 2006. *In vitro* studies of the digestion of caprine whey proteins by human gastric and duodenal juice and the effects on selected microorganisms. The British Journal of Nutrition 96, 562–9.

Anema, S., Stanley, D., 1998. Heat-induced, pH-dependent behaviour of protein in caprine milk. International Dairy Journal 8, 917–923.

Ceballos, L., Morales, E., Martínez, L., Extremera, F., Sampelayo, M., 2009. Utilization of nitrogen and energy from diets containing protein and fat derived from either goat milk or cow milk. Journal of Dairy Research 76, 497–504.

De Jong, P., 1997. Impact and control of fouling in milk processing. Trends in Food Science & Technology 8, 401–405.

Haenlein, G., 2004. Goat milk in human nutrition. Small Ruminant Research 51, 155–163.

Hurley, W.L., Theil, P.K., 2011. Perspectives on immunoglobulins in colostrum and milk. Nutrients 3, 442–74.

Jandal, J., 1996. Comparative aspects of goat and sheep milk. Small Ruminant Research 22, 177–185.

Kapila, R., Kavadi, P., Kapila, S., 2013. Comparative evaluation of allergic sensitization to milk proteins of cow, buffalo and goat. Small Ruminant Research 112, 191–198.

Kontopidis, G, Holt, C, Sawyer, L, 2004. Invited review: β-lactoglobulin: binding properties, structure, and function. Journal of Dairy Science 87, 785–796.

Lara-Villoslada, F., Olivares, M., Jiménez, J., Boza, J., Xaus, J., 2004. Goat milk is less immunogenic than cow milk in a murine model of atopy. Journal of Pediatric Gastroenterology and Nutrition 39, 354–360.

Pesic, M., Barac, M., Stanojevic, S., 2016. Heat-Induced Casein–Whey Protein Interactions in Caprine Milk: Whether Are Similar to Bovine Milk? Springer.

Saarinen, K.M., Pelkonen, A.S., Mäkelä, M.J., Savilahti, E., 2005. Clinical course and prognosis of cow's milk allergy are dependent on milk-specific IgE status. The Journal of Allergy and Clinical Immunology 116, 869–75.

Zadow, J.G., Hardham, J.F., Kocak, H.R., Mayes, J.J., 1983. The stability of goat's milk to UHT processing. The Australian Journal of Dairy Technology 1, 20–23.

8. Appendix

Table 8.1: Protein Data Bank ID's referenced within this thesis.

PDB ID	Species	Variant	pH	Space Group	Reference
1BSO	Cow	A	7.3	$P3_221$	Qin *et al.,* 1998b
1QG5	Cow	A	7.9	$C222_1$	Oliveira *et al.*, 2001
1B8E	Cow	B	7.9	$C222_1$	Oliveira *et al.*, 2001
2R56	Cow	B	5.5	$P2_12_12_1$	Niemi *et al.* 2007
1BEB	Cow	A/B	7.8	P1	Brownlow *et al.*, 1997
3BLG	Cow	A	6.2	$P3_221$	Qin *et al.*, 1998a
1BSY	Cow	A	7.1	$P3_221$	Qin *et al.*, 1998a
2BLG	Cow	A	8.2	$P3_221$	Qin *et al.*, 1998a
1BSQ	Cow	B	7.1	$P3_221$	Qin *et al.*, 1999
4TLJ	Goat		6.8	$P12_11$	Crowther *et al.,* 2014
4OMW	Goat		7.5	$P2_12_12$	Loch *et al.* 2015
4OMX	Goat		8.5	$P3_121$	Loch *et al.* 2015
4CK4	Ovine		5.6	P1	Kontopidis *et al.*, 2014
4NLI	Ovine		7.0	$P3_121$	Loch *et al.*, 2014
4NLJ	Ovine		7.0	P1	Loch *et al.*, 2014
1YUP	Reindeer		6.5	P1	Oksanen *et al.,* 2006
1EXS	Porcine		3.2	$P3_221$	Hoedemaeker *et al.*, 2002

References:

Brownlow, S., Cabral, J., Cooper, R., Flower, D., Yewdall, S., Polikarpov, I., North, A., Sawyer, L., 1997. Bovine β-lactoglobulin at 1.8 Å resolution — still an enigmatic lipocalin. Structure 5, 481–495.

Crowther, J., Lassé, M., Suzuki, H., Kessans, S., Loo, T., Norris, G., Hodgkinson, A., Jameson, G., Dobson, R., 2014. Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin. FEBS Letters 588, 3816–3822.

Hoedemaeker, F.J., Visschers, R.W., Alting, A.C., de, K.G., Kuil, M.E., Abrahams, J.P., 2002. A novel pH-dependent dimerization motif in β--lactoglobulin from pig (Sus scrofa). Acta Crystallogr Sect D Biological Crystallogr 58, 480–486.

Kontopidis, G., Gilliver, A., Sawyer, L., 2014. Ovine β-lactoglobulin at atomic resolution. Acta Crystallographica Section F: Structural Biology Communications 70, 1498–1503.

Loch, J., Bonarek, P., Polit, A., Świątek, S., Czub, M., Ludwikowska, M., Lewiński, K., 2015. Conformational variability of goat β-lactoglobulin: Crystallographic and thermodynamic studies. International Journal of Biological Macromolecules 72, 1283–1291.

Loch, J., Molenda, M., Kopeć, M., Świątek, S., Lewiński, K., 2014. Structure of two crystal forms of sheep β-lactoglobulin with EF-loop in closed conformation. Biopolymers 101, 886–894.

Niemi, M., Jylhä, S., Laukkanen, M.-L., Söderlund, H., Mäkinen-Kiljunen, S., Kallio, J., Hakulinen, N., Haahtela, T., Takkinen, K., Rouvinen, J., 2007. Molecular Interactions between a Recombinant IgE Antibody and the β-Lactoglobulin Allergen. Structure 15, 1413–1421.

Oksanen, E., Jaakola, V.P., Tolonen, T., Valkonen, K., Åkerström, B., Kalkkinen, N., Virtanen, V., Goldman, A., 2006. Reindeer β-lactoglobulin crystal structure with pseudo-body-centred noncrystallographic symmetry. Acta Crystallographica Section D: Biological Crystallography 62, 1369–1374.

Oliveira, K.M., Valente-Mesquita, V.L., Botelho, M.M., Sawyer, L., Ferreira, S.T., Polikarpov, I., 2001. Crystal structures of bovine β-lactoglobulin in the orthorhombic space group C2221. European Journal of Biochemistry 268, 477–484.

Qin, B., Bewley, M., Creamer, L., Baker, H., Baker, E., Jameson, G., 1998a. Structural Basis of the Tanford Transition of Bovine β-Lactoglobulin. Biochemistry 37, 14014–14023.

Qin, B., Creamer, L., Baker, E., Jameson, G., 1998b. 12-Bromododecanoic acid binds inside the calyx of bovine β-lactoglobulin. FEBS Letters 438, 272–278.

Qin, B., Jameson, G., Bewley, M., Baker, E., Creamer, L., 1999. Functional implications of structural differences between variants A and B of bovine β-lactoglobulin. Protein Science 8, 75–83.

Publications included in this thesis:

Crowther, J.M., Jameson, G.B., Hodgkinson, A.J., and Dobson, R.C.J., (2016). Structure, Oligomerisation and Interactions of β-Lactoglobulin, Milk Proteins - From Structure to Biological Properties and Health Aspects, Dr. Isabel Gigli (Ed.), InTech, DOI: 10.5772/62992.

Crowther, J.M., Lassé, M., Suzuki, H., Kessans, S.A., Loo, T.S., Norris, G.E., Hodgkinson, A.J., Jameson, G.B. and Dobson, R.C.J. (2014), Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin, FEBS Letters, 588, doi: 10.1016/j.febslet.2014.09.010.

Chapter 3

Structure, Oligomerisation and Interactions of β-Lactoglobulin

Jennifer M. Crowther, Geoffrey B. Jameson,
Alison J. Hodgkinson and Renwick C.J. Dobson

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/62992

Abstract

β-Lactoglobulin (βLg), as the most abundant whey protein in ruminant milk and as a useful model protein, is the subject of countless biophysical studies in the literature, yet its physiological role is hitherto unknown. This chapter deals with studies that focus on the structure of βLg, its oligomeric behaviour and the interactions that this protein participates in. These and further studies are necessary to understand how the protein's physicochemical properties may influence the processing, digestion and immunogenicity of ruminant milks and their products. However, there is also a need for research into the interactions that occur naturally between βLg and other components in milk, as this may give us insight into the physiological role of the protein.

Keywords: β-lactoglobulin, milk protein, lipocalin, oligomeric state, structure, Tanford transition

1. Introduction

β-Lactoglobulin (βLg) is found in the milk of many mammals and, although it is the most abundant protein present in the whey fraction of ruminant milk [1], it is completely absent in human milk. Due to its abundance and relative ease of purification, bovine βLg has served as a model protein for countless biophysical studies of folding, stability and self-association. Although this has created an extensive literature on the nature of βLg, its physiological function is yet to be determined.

βLg belongs to the lipocalin family of proteins, most of which have roles that involve ligand-binding [2]. Its ability to bind hydrophobic molecules *in vitro* has prompted speculation that βLg is involved in the transport of insoluble and/or chemically sensitive molecules between mother and offspring. However, it is necessary to draw the distinction between demonstrating binding *in vitro* and identifying an endogenous ligand that translates to a physiological role of βLg *in vivo*.

Understanding the behaviour of this protein is of particular interest to the dairy industry, given the potential of βLg to affect the processing and manufacture of milk products; for example, βLg aggregation upon heat treatment is known to contribute to the fouling of heat exchangers during the processing of milk [3]. βLg has also been identified as one of the main immunogenic proteins in cow milk and thus contributes to cow milk allergies [4]. It is, therefore, of value to understand this protein's physicochemical properties and how they may influence the processing, digestion and immunogenicity of ruminant milk and their products.

The purpose of this chapter is to review the knowledge that has been gathered for a range of βLg orthologues from various species with regard to structure, oligomerisation and interaction behaviour under predominantly physiological conditions and to consider the current gaps in our knowledge. The thermal denaturation behaviour of βLg, including heat-induced interactions and fibril formation, has been dealt with in detail elsewhere [5].

2. Structure of βLg

The first reported atomic level resolution structure of βLg, solved by X-ray crystallography for bovine βLg [6], showed remarkable similarity to retinol-binding protein and led to the classification of βLg as a lipocalin. Lipocalins are a family of proteins that share a similar structure despite great diversity at the sequence level. The conserved lipocalin fold comprises an eight-stranded anti-parallel β-sheet (strands A–H) that is folded back upon itself to enclose an internal cavity, often termed a calyx, together with a three-turn α-helix calyx (cup) handle that lies approximately above strand H (**Figure 1**) [7]. This fold allows lipocalins to bind a wide range of hydrophobic ligands, with the cavity size and loop scaffold at the cavity entrance determining selectivity. While they were once simply classified as transport proteins, lipocalins are now known to exhibit vast functional diversity, yet most involve some form of ligand binding [2].

Numerous high-resolution atomic structures now exist for bovine βLg, along with structures of orthologues from sheep, goat and reindeer and the more distantly related pig. Like other lipocalins, bovine βLg is a small protein, in this case of 162 amino acids with a monomeric mass of ~18,300 Da. As shown in a three-dimensional cartoon form in **Figure 2**, each subunit consists predominantly of an antiparallel β-sheet formed by eight β-strands, A–H, wrapped around to form a flattened calyx [9]. The calyx is flanked on its outer surface by a three-turn alpha helix. The dimer interface, at least for βLg from ruminants, is formed by the ninth β-strand (I) along with the loop connecting strands A and B. The loops at the closed end of the

http://dx.doi.org/10.5772/62992

calyx (BC, DE and FG) are quite short, whereas those at the open end (AB, CD, EF and GH) are longer and more flexible [10].

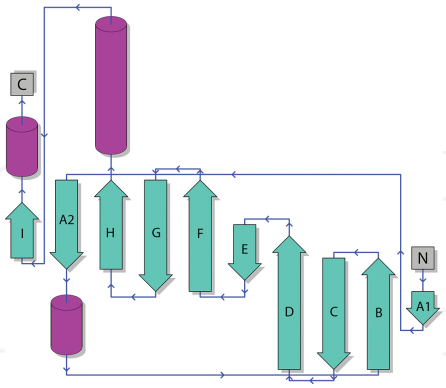


Figure 1. Topology diagram showing characteristic features of lipocalin proteins. The eight β-strands A–H form the antiparallel β-barrel. Strand A is kinked (giving rise to strands A1 and A2) and connects the β-sheet comprising strands A1, B, C and D to the β-sheet comprising strands E, F, G, H and A2. The commonly occurring N-terminal 3_{10} helix and the ubiquitous 3-turn α-helix following strand H are shown. Generally, a ninth β-strand follows the 3-turn α-helix and is packed against strand A2. Figure generated by PDBsum [8] using the structure of bovine βLg, PDB ID: 1BSO.

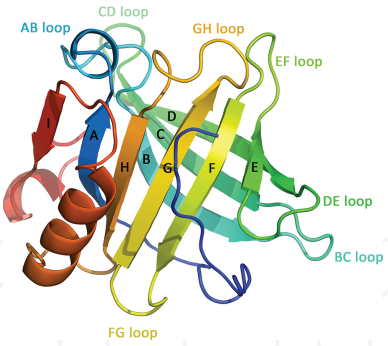


Figure 2. Crystal structure of one monomer of bovine βLg (PDB ID: 1BSO). Eight β-strands (A–H) form the central antiparallel β-sheet calyx. The calyx is flanked on its outer surface by a three-turn α-helix. The ninth β-strand, I, and the AB loop are involved in dimer formation. The polypeptide chain is shown in rainbow colours, beginning in blue at the N-terminus and ending in red at the C-terminus.

Each monomer of βLg contains five cysteine residues. One exists as a free thiol on strand G and is buried beneath the α-helix that lies alongside the β-barrel, whereas the other four form two disulfide bridges. The first (Cys66-Cys160) links the C-terminus to the CD loop, while the second (Cys106-Cys119) links β-strands G and H [11]. The presence and correct arrangement of these disulfide bonds are crucial for the correct folding of βLg. Recombinant expression of

this protein in a soluble form in bacteria requires the simultaneous coexpression of a disulfide bond isomerase along with the use of modified *Escherichia coli* Origami (DE3) cells [12]. These cells, which carry mutations in the thioredoxin reductase and glutathione reductase genes, provide an oxidising environment and, together with the disulfide bond isomerase, allow the proper formation of disulfide bonds in the cytoplasm [13].

At least eleven protein sequence variants of bovine βLg have been described, with A and B the most common variants [14]. Variant B differs from A by two amino acid substitutions: Asp64Gly in the mobile surface loop (CD) and Val118Ala in the hydrophobic core [15]. Crystal structures of both variants A and B at pH 7.1 have allowed observation of the structural consequences of these sequence differences [11]. Only minor differences can be seen in the calyx, while small changes occur in the main chain conformations in the vicinity of the Asp64Gly mutation. The substitution of Val118 for Ala causes changes in the local structure creating a void volume that weakens several hydrophobic interactions. This may be responsible for the slight decrease in thermal stability of variant B relative to variant A [11]. The conformation of the EF loop is slightly different, but this may be due to differences in crystallisation or imprecision in definition due to the high mobility of this loop.

Crystal structures of ovine (sheep) [16, 17], caprine (goat) [18, 19] and reindeer βLg [20] indicate that these orthologues share a high degree of structural similarity with bovine βLg, at both the tertiary and quaternary level (**Figure 3**), with minimal root-mean-square deviations when aligning the C-α atoms of these structures (**Table 1**). However, there are significant differences between the structures of these orthologues and that of porcine βLg, which in the crystal structure features a completely different quaternary association [21]. This is not unreasonable considering the lower level of sequence identity (63–65%) between porcine βLg and bovine, ovine, caprine and reindeer βLg, where the latter four share sequence identity in pairwise comparisons of 93–99% (**Table 1**).

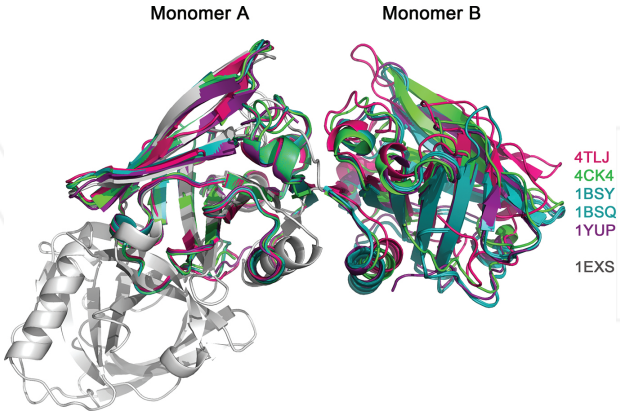


Figure 3. Crystal structure of caprine βLg (PDB: 4TLJ, pink) overlaid with bovine βLg structures (PDB: 1BSY, 1BSQ, blue), ovine βLg (PDB: 4CK4, green), reindeer βLg (PDB: 1YUP, purple) and porcine βLg (PDB: 1EXS, grey). Monomer A of each structure is superposed to highlight the variation in orientation of the second monomer.

	Bovine A	Bovine B	Ovine	Caprine	Reindeer	Porcine
Bovine A						
% sequence identity	100	98.77	95.06	95.06	93.21	65.00
rmsd monomer	-	0.208	0.463	0.397	0.429	1.849
rmsd dimer	-	0.221	1.023	1.400	0.569	19.36
Bovine B						
% sequence identity	98.77	100	96.30	96.30	94.44	63.75
rmsd monomer	0.208	-	0.409	0.386	0.394	1.792
rmsd dimer	0.221	-	0.964	1.318	0.496	19.13
Ovine						
% sequence identity	95.06	96.30	100	99.38	93.21	62.50
rmsd monomer	0.463	0.409	-	0.324	0.310	1.920
rmsd dimer	1.023	0.964	-	1.076	0.739	19.49
Caprine						
% sequence identity	95.06	96.30	99.38	100	93.21	62.50
rmsd monomer	0.397	0.386	0.324	-	0.404	1.398
rmsd dimer	1.400	1.318	1.076	-	1.242	19.69
Reindeer						
% sequence identity	93.21	94.44	93.21	93.21	100	65.00
rmsd monomer	0.429	0.394	0.310	0.404	-	1.723
rmsd dimer	0.569	0.496	0.739	1.242	-	19.06
Porcine						
% sequence identity	65.00	63.75	62.50	62.50	65.00	100
rmsd monomer	1.849	1.792	1.920	1.398	1.723	-
rmsd dimer	19.36	19.13	19.49	19.69	19.06	-

Source: PDB IDs for structures used for alignment, performed in PyMol: Bovine A: 1BSY, Bovine B: 1BSQ, Ovine: 4CK4, Caprine: 4TLJ, Reindeer: 1YUP and Porcine: 1EXS.

Table 1. The percentage sequence identity of βLg orthologue protein sequences, the rmsd when aligning the Cα atoms of monomers and the rmsd when aligning the Cα atoms of dimers.

Although there is a high level of structural similarity among the bovine, caprine, ovine and reindeer orthologues, there are subtle differences between them. When dimers are selected for alignment the root-mean-square deviations for the superposition of these structures are higher than when a single monomer is used for the alignment (**Table 1**). This indicates that while the tertiary fold of these orthologues is similar, there is flexibility in the orientation of the monomers relative to each other (**Figure 3**), with different crystal forms, including those of bovine βLg, sampling different conformations.

The ultrahigh resolution crystal structures of caprine and ovine βLg [17, 18] make it possible to clearly define features that in lower resolution bovine and reindeer βLg structures are

obscured by disorder and conformational promiscuity. These features include the long flexible CD and GH loops, the C-terminal region, and the AB loops at the dimer interface. However, the more mobile regions of bovine and reindeer βLg structure loops are also the more mobile regions of the caprine and ovine βLg structures. The dimer interface in these ruminant βLg structures is formed by the antiparallel association of the I β-strands and electrostatic interactions between Asp33 and Arg40 residues located within the AB loop of each monomer (**Figure 4**). Hydrogen bonding between the main chain of Ala34 and the side chains of the Asp/Arg pair holds this residue (Ala34) in an unfavourable conformation. The only other residue found in a less favourable region of the Ramachandran plot is the highly conserved Tyr99, which forms part of a γ-turn [22].

Bovine βLg undergoes several conformational changes with pH. Several studies have used nuclear magnetic resonance (NMR) spectroscopy to determine the three-dimensional structure of bovine βLg at pH 2.6 [23–25], where at very low ionic strength the protein becomes monomeric (see the next section for more details on the oligomeric behaviour of βLg). While monomer–dimer exchange makes it difficult to obtain NMR structures at neutral pH [26], this

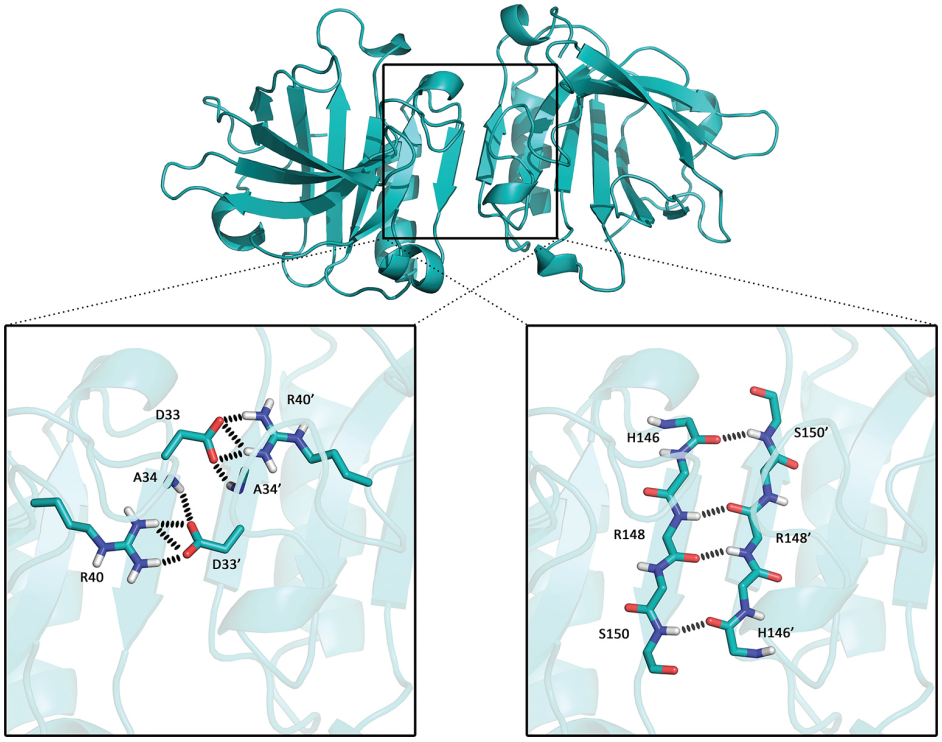


Figure 4. Dimer interface of bovine βLg [PDB ID: 1BEB]. Close-up views of AB-loops and I-strands are shown where inter subunit hydrogen bonds and electrostatic interactions between side chains and main chains can be seen.

http://dx.doi.org/10.5772/62992

low pH NMR structure can be compared to crystal structures of bovine βLg solved at higher pH values. The structure of this low pH form is very similar to a subunit of the dimer at pH 6.2 [23]. There are slight deviations in the orientation of the loops and of the three-turn helix flanking the calyx, but overall the tertiary fold and, in particular, the hydrophobic cavity are well maintained at low pH.

Bovine βLg undergoes another reversible pH-induced conformational change, termed the Tanford transition [26, 27], at close to physiological pH that may be of functional significance. This transition involves movement of the EF loop, which is located at the mouth of the calyx. The EF loop is in a closed conformation at pH ~7.1 and below [11] , burying Glu89 as glutamic acid, whereas it adopts an open conformation at pH values above 7, exposing Glu89 as a glutamate. The structures of ovine, caprine and reindeer βLg solved at pH 6.5–7, 6.8 and 6.9, respectively, all show the EF loop in the closed position, in agreement with bovine structures below pH 7. A recent structure of caprine βLg [19] shows an asymmetric dimer with the EF loop of one subunit in the closed position and the other in the open position, suggesting that goat βLg also undergoes the Tanford transition. It is possible that the Tanford transition plays a role in regulating the ligand-binding properties of βLg.

3. Oligomerisation

The oligomerisation of βLg has been studied intensively using various techniques including analytical ultracentrifugation, isothermal titration calorimetry and small angle X-ray scattering [18, 28–31]. Utilising both sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation, Mercadante et al. [31] investigated the oligomerisation of bovine βLg over a pH range of 2.5–7.5. Interestingly, at pH 2.5, 3.5, 6.5 and 7.5, the weight-averaged sedimentation coefficient increased with increasing protein concentration, suggesting a monomer–dimer equilibrium. However, at pH 4.5 and 5.5, the weight-averaged sedimentation coefficient stayed the same, consistent with a single species (a dimer) predominating across the concentrations used in the experiment. This suggests that the dimer is more strongly associated at pH 4.5 and 5.5, values which lie near the isoelectric point of the protein, than at pH 2.5, 3.5, 6.5 and 7.5.

Mercadante et al. [31] went on to characterise the binding energy of dimer formation as a function of ionic strength. They demonstrated that an increase in ionic strength strongly favours formation of the dimer. At low pH, dissociation of the dimer is extremely sensitive to ionic strength. This is due to the fact that at low pH, ionic strength stabilises the dimer by the association of anions near the dimer interface which mitigate charge repulsion of the positively charged subunits. On the other hand, at neutral pH, a relatively low density of cations in the region of the AB and GH loops can also help to stabilise the dimer, but the effect of ionic strength is less pronounced due to the smaller magnitude of the nominal charge on the protein (−9 at pH ~7.5 and +20 at pH ~2.5).

A recent paper [32], featuring the use of synchrotron FT-IR techniques to study the dimer–monomer equilibrium at pH ~7 of bovine βLg at high salt and high protein concentrations, reported dimer dissociation constants orders of magnitude greater that the micro molar values

reported by a host of others (see Supplementary Table in reference (31). However, inspection of the FT-IR data would support an interpretation that at the higher concentrations aggregation and denaturation of originally dimeric βLg were occurring.

Importantly, the study by Mercadante et al. [31] indicates that under physiological conditions bovine βLg self-associates into a dimer with a dissociation constant within the range of the concentrations studied (5–45 μM). Bovine βLg is, therefore, likely to be dimeric at the concentration and pH typically associated with milk (~3 mg mL^{-1} or ~165 μM and pH 6.5 [33]). A study of caprine βLg suggests a comparable dissociation constant as for bovine βLg under similar conditions [18]. Given the high level of sequence identity between caprine and ovine βLg [17], it is likely that ovine βLg exhibits similar oligomerisation behaviour. The nonruminant equine and porcine βLg orthologues, however, are monomeric at physiological pH [21, 34]. In contrast to bovine βLg, through a domain-swapping mechanism, porcine βLg forms a dimer at low pH [21].

The dimer interface of bovine, caprine, ovine and reindeer βLg orthologues is composed of an intermolecular β-sheet formed between the I-strands of each monomer along with electrostatic interactions and hydrogen bonds between residues located on the AB loops (see **Figure 4**). Sakurai and Goto [28] investigated the impact of these elements on the monomer–dimer equilibrium of bovine βLg at neutral pH using analytical ultracentrifugation. They engineered mutants that either disrupted the intermolecular β-sheet or the electrostatic interactions between the AB loops. All of the mutants that introduced a proline residue within the I-strand led to dissociation of the dimer, due to disruption of the intermolecular β-sheet. Substitution of the Asp33 or Arg40 residues of the AB loop with oppositely charged residues was also unfavourable for dimer formation, due to the electrostatic repulsion introduced. When these charge mutants were mixed they were able to form heterodimers, suggesting that the electrostatic interactions between these aspartate and arginine residues contribute to stabilisation of the dimer.

Some orthologues, however, such as equine βLg, exist as monomers at neutral pH despite the presence of the I-strands and AB loop residues [35]. Sakurai and Goto [28] created another set of bovine βLg mutants in which they substituted the remaining residues at the dimer interface with those found in the monomeric equine and porcine βLg sequences. These sequence-based mutations did not largely affect the association constant, indicating that dimer stabilisation cannot be ascribed simply to the interface residues of bovine βLg that differ in the monomeric βLg variants.

In a similar, but opposite, experiment Kobayashi et al. [36] aimed to convert the monomeric equine βLg to a dimeric form by substituting I-strand and AB loop residues with those found in bovine βLg. That is, Ser34 and Glu35 in the AB loop of equine βLg were replaced with Ala and Gln, and the sequence comprising the I-strand was replaced with the corresponding bovine βLg sequence. Interestingly these mutants did not form a dimer, further suggesting that the difference in oligomerisation behaviour between the bovine and equine orthologues cannot be explained simply by the sequence differences at the dimer interface. They hypothesised that structural differences must exist between equine and bovine βLg that prevent the same interactions occurring at the dimer interface.

http://dx.doi.org/10.5772/62992

To assess this, Ohtomo et al. [35] constructed a chimera named Gyuba, which means cow and horse in Japanese. Gyuba was made by joining the secondary structural elements of bovine βLg with the loops of equine βLg. The chimera was able to form dimers, and its crystal structure showed that it had a very similar dimer interface as seen for bovine βLg (PDB ID: 3KZA). Taken together, these studies suggest that the entire arrangement of the secondary structural elements and loops of βLg, including hydrophobic interactions, hydrogen bonds between I-strands, and electrostatic interactions and hydrogen bonds at the AB loops, is necessary for dimerisation. Further, it is tempting to speculate that protein dynamics may also play a role in dimer formation.

4. Interactions

Due to its similarity to retinol-binding protein, the ability of βLg to bind retinol was examined by fluorescence spectroscopy [34]. The fact that retinol was able to bind to βLg, and that it could be modelled into the crystal structure of βLg in a similar position as seen bound in retinol-binding protein, led to speculation that the biological function of βLg is to transport vitamin A in milk [6]. However, since then bovine βLg has been shown to be capable of binding a range of small hydrophobic molecules, as demonstrated in various ligand-bound crystal structures. These bound ligands include vitamin D [33, 37], vitamin A [38], cholesterol [33], a range of fatty acids [39–43] and the fatty-acid derivative 12-bromododecanoic acid [9], as well as more diverse molecules including SDS [44] and various anaesthetic drugs [45]. Ligand binding has also been investigated by a variety of other methods, including equilibrium dialysis, affinity chromatography, electron spin resonance spectroscopy, spectrophotometry and perturbation of intrinsic tryptophan fluorescence [33, 46]. Provided that there is a chromophore on the ligand, induced circular dichroism provides, along with X-ray crystallography, the most definitive method for characterising the binding of ligands. These studies have revealed a broad range of ligands that βLg is capable of binding to.

Ligand-bound crystal structures serve as definitive proof that small hydrophobic ligands are accommodated within the hydrophobic calyx of βLg. Ligand binding, therefore, is dependent on the opening of the EF loop at the mouth of the calyx which, for bovine βLg, occurs near pH 7 as Glu89 becomes deprotonated and is exposed as a glutamate. The pKa of Glu89 in porcine βLg is higher than in bovine βLg (calculated 9.7 compared to observed 7.3 (see reference [11] and references therein to Tanford's original work)), and thus porcine βLg is only able to bind fatty acids above pH 8.6 [47]. In ligand-bound structures the density is often quite poor for the extremity of the ligand and can be ambiguous as to which orientation the molecule faces. However, most can be interpreted by taking into account their chemistry, for instance cholesterol and vitamin D were built into their electron density placing their hydroxyl groups facing out of the calyx as opposed to being buried in the centre of the protein [33]. Qin et al. [9] used a fatty acid derivative, 12-bromododecanoic acid, to unequivocally determine the orientation of the ligand in the binding site, with the carboxylate head group lying at the surface of the molecule.

The lining of the hydrophobic cavity is exclusively hydrophobic, except for two lysine residues (Lys60 and Lys69) at the entrance to the calyx. It is generally agreed that there are two main interactions between βLg and fatty acid ligands; one is the hydrophobic interaction between the hydrocarbon tail of the fatty acid and the interior of the hydrophobic calyx, and the other is the electrostatic attraction between the carboxyl group of the fatty acid and the amines of Lys60 and Lys69. The strength of the interaction between βLg and fatty acids generally increases as the length of the hydrocarbon chain increases, due to an increase in van der Waals' forces. There is, however, an exception to this rule; the eight-carbon caprylic acid has a greater binding affinity than the ten-carbon capric acid [43]. Until recently, there was no satisfactory explanation for this result. Yi and Wambo [48] have used molecular dynamics simulations to accurately compute the binding free energies between βLg and five saturated fatty acids of 8 to 16 carbon atoms. Their results agree well with experimental results; the binding free energy increases with the number of carbon atoms of the fatty acid, with the exception of caprylic acid, which has a higher binding free energy than the 2 carbon longer capric acid. The van der Waals' forces between the fatty acid tails and the interior of the βLg calyx increase as the chain length increases; however, for caprylic acid the electrostatic interaction between the carboxyl group and the amines of Lys60 and Lys69 is stronger than these van der Waals' forces which pulls the caprylic acid closer to the top of the barrel. This allows the hydrophobic tail of caprylic acid to fluctuate more easily, increasing the entropy of this complex, resulting in a greater overall binding free energy.

A small number of studies suggests there may be a second, lower affinity, external binding site for hydrophobic molecules [37, 49, 50]. The lower affinity of this binding site may explain the difficulty in obtaining crystallographic evidence of this interaction. However, a crystal structure was recently solved of bovine βLg that identifies two molecules of vitamin D_3 bound, one bound within the calyx and the second possibly bound at an exosite between the β-barrel and the α-helix that lies alongside the barrel (**Figure 5**) [37]. The free thiol of βLg (Cys121) is buried beneath this α-helix and methylation of this thiol appears to reduce the affinity for palmitic acid compared to native βLg, whereas the binding of retinol is not affected [49]. A second binding site may explain how βLg is capable of binding such a wide diversity of shapes of ligands, yet more evidence, such as more convincing electron density, is required before the existence of this site can be conclusively proven.

A definitive role for βLg is yet to be ascribed, although several predictions have been made. Most suggestions are for a role in molecular transport between mother and offspring, due to its demonstrated ability to bind a range of ligands. However, the specific identity of the ligand being transported is not clear. Fatty acids have been seen bound to βLg isolated from milk under nondenaturing conditions, but are present in quantities reflecting the fatty acid composition of milk [51]. The apparent lack of selectivity makes it less likely that βLg is a specific fatty acid or vitamin transporter. βLg may still be involved in fat metabolism; there is evidence that βLg can stimulate a pregastric lipase, potentially by binding the fatty acid products and thereby reducing their inhibitory effect on the enzyme [52].

http://dx.doi.org/10.5772/62992

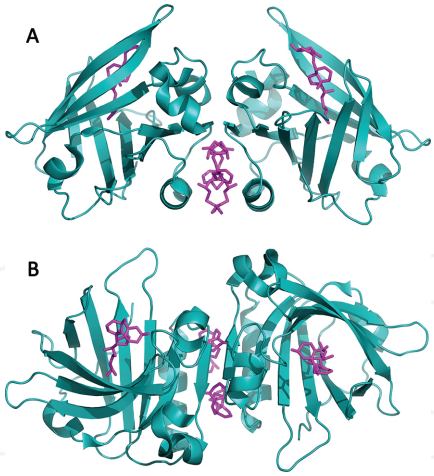


Figure 5. Front view (A) and top view (B) of the crystal structure of bovine βLg showing two molecules of vitamin D_3 (magenta) bound, indicating the locations of the putative primary ligand-binding site within the central hydrophobic calyx and the second proposed exosite between the β-barrel and the α-helix that lies alongside the barrel. (PDB ID: 2GJ5 [37]).

The closest homologue to βLg in humans is glycodelin (pregnancy protein 14). Inter alia, glycodelin has an immunosuppressive activity in the uterus and is involved in protecting products of the reproductive organs from the immune system [53]. Unlike βLg, glycodelin is a glycoprotein, a property essential for its function. It has been hypothesised that βLg may have diverged from glycodelin following a gene duplication event and has since lost all glycodelin-related function [33]. βLg may now exist primarily as an important source of amino acids for the offspring of the animals that produce it. However, the resistance of this protein to proteolysis by pepsin [54] along with the high level of sequence conservation seen among βLg orthologues, including the highly conserved Glu89 within the EF loop, argue against a simple nutritive function.

Another enticing proposal is the notion that βLg may possess antimicrobial activity [55–57]. The intact protein appears to be capable of inhibiting the growth of *Staphylococcus aureus*, *Streptococcus uberis* and *E. coli* bacteria largely responsible for the prevalent and costly disease, bovine mastitis [55]. βLg also appears to augment the antimicrobial activity of lactoferrin, a protein with a known role in the defence against mammary gland infections [56]. Pan et al. [58], however, reported a lack of antimicrobial activity of native βLg, yet this may be due to the use of a commercial preparation of βLg, as compared to the mild, nondenaturing isolation from milk employed in the aforementioned studies. Pellegrini et al. [57] have also described the antimicrobial activity of four peptides derived from βLg following trypsin digestion. This may point to a protective physiological role in new born calves, perhaps in addition to a protective role in the secretory tissue of the mother. Further studies are needed to provide a comprehensive understanding of the relevance of these findings and to address the mechanisms underlying these antimicrobial observations.

Upon heating, βLg (both bovine and caprine) associates with casein micelles through formation of a βLg/casein complex [59], with covalent intermolecular disulfide bonds forming above 75 °C [60]. The pH at which the milk is heated is important in determining how much complex is formed as well as the extent of dissociation of casein from the micelles, which is implicated in the heat stability of the milk [59, 60]. Further studies such as these are necessary for understanding how the processing of both bovine and nonbovine milk may affect the properties of milk constituents and how these may then influence the digestion and immunogenicity of ruminant milk and their products.

Much research has been devoted to understanding the functionality of βLg in milk. There remain large gaps in our knowledge of the interactions of proteins in milk under physiological conditions. There is, therefore, a critical need for research into the interactions that occur between βLg and protein components (other than caseins) in milk under physiological/untreated conditions, as this may finally give us insight into the actual physiological role of this protein and to identify factors that distinguish human neonate responses to milk products sourced from different ruminants. In this regard, the structure of bovine βLg with human immunoglobulin fragments is highly significant [61].

5. Conclusion

βLg has served as the focus of an extensive range of studies for well over half a century, creating a wealth of knowledge about this enigmatic protein. We now have a clear view of the native structure of βLg. This small globular protein is characterised by a central β-barrel composed of eight antiparallel β-strands, creating a calyx that is well suited to binding hydrophobic ligands. An α-helix lies alongside the barrel potentially creating a second, lower-affinity, binding site for ligands. A ninth β-strand, along with the loop connecting the A and B strands, forms the dimer interface of ruminant βLg.

Understanding the structure of βLg has given considerable insight into its behaviour in solution. Under physiological conditions, ruminant βLg orthologues are predominantly dimeric. The dimer interface involves 12 intermolecular hydrogen bonds and 2 ion pairs, interactions that have been shown to be critical for dimer stability. At low pH the protein is positively charged and thus under low-salt conditions it is monomeric. Increasing the ionic strength screens these electrostatic repulsions and stabilises dimer formation.

Many questions regarding βLg remain and, in particular, the physiological function of the protein is still a mystery. The proven ability of βLg to bind hydrophobic molecules along with its stability at low pH, and resistance to proteolytic enzymes are strongly suggestive of a role in fatty acid transport between mother and child. Alternatively, these qualities may enable βLg to enhance milk fat metabolism through the promotion of pregastric lipase activity. However, the absence of βLg in other species, most notably human, needs to be remembered when considering the role of βLg.

The physicochemical properties of bovine βLg will, undoubtedly, continue to be investigated. What is required now is a detailed understanding of these properties in closely related

orthologues, in order to understand the underlying processes occurring during the processing and digestion of different ruminant milks. There is also a significant need for exploration into the interactions with βLg that are occurring naturally in milk. This may provide the necessary insight into the function of this protein that is of physiological significance to the mother and/or her offspring, and into the functionality of this protein in milk products from different ruminants destined for human consumption.

Author details

Jennifer M. Crowther[1], Geoffrey B. Jameson[2], Alison J. Hodgkinson[3] and Renwick C.J. Dobson[1,4*]

*Address all correspondence to: renwick.dobson@canterbury.ac.nz

1 Biomolecular Interaction Centre, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

2 Institute of Fundamental Sciences and the Riddet Institute, Massey University, Palmerston North, New Zealand

3 Food and Bio-based Products, AgResearch Limited, Ruakura Research Centre, Hamilton, New Zealand

4 Department of Biochemistry and Molecular Biology, Bio21 Institute, University of Melbourne, Parkville, Victoria, Australia

References

[1] Hinz K, O'Connor P, Huppertz T, Ross R, Kelly A. Comparison of the principal proteins in bovine, caprine, buffalo, equine and camel milk. The Journal of Dairy Research 2012;79:185–191.

[2] Sawyer L, Kontopidis G. The core lipocalin, bovine β-lactoglobulin. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 2000;1482:136–148.

[3] Galani D, Apenten R. Heat-induced denaturation and aggregation of β-lactoglobulin: kinetics of formation of hydrophobic and disulphide-linked aggregates. International Journal of Food Science & Technology 1999;34:467–476.

[4] Kapila R, Kavadi P, Kapila S. Comparative evaluation of allergic sensitization to milk proteins of cow, buffalo and goat. Small Ruminant Research 2013;112:191–198.

[5] Edwards PJB, Jameson GB. 2014. Structure and stability of whey proteins. In Milk Proteins (2nd edition): From Expression to Food. Boland M, Singh H, Thompson A, Eds. Elsevier, San Diego. Chapter 7, pp 201–242.

[6] Papiz MZ, Sawyer L, Eliopoulos EE, North AC, Findlay JB, Sivaprasadarao R, et al. The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein. Letters to Nature 1986;324:383–385.

[7] Flower DR. The lipocalin protein family: structure and function. Biochemical Journal 1996;318:1–14.

[8] Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM. PDBsum: a web-based database of summaries and analyses of all PDB structures. Trends in Biochemical Sciences 1997;22:488–490.

[9] Qin BY, Creamer LK, Baker EN, Jameson GB. 12-Bromododecanoic acid binds inside the calyx of bovine β-lactoglobulin. FEBS Letters 1998;438:272–278.

[10] Jameson GB, Adams JJ, Creamer LK. Flexibility, functionality and hydrophobicity of bovine β-lactoglobulin. International Dairy Journal 2002;12:319–329.

[11] Qin BY, Jameson GB, Bewley M, Baker EN, Creamer LK. Functional implications of structural differences between variants A and B of bovine β-lactoglobulin. Protein Science 1999;8:75–83.

[12] Ponniah K, Loo TS, Edwards PJB, Pascal SM, Jameson GJ, Norris GE. The production of soluble and correctly folded recombinant bovine β-lactoglobulin variants A and B in Escherichia coli for NMR studies. Protein Expression and Purification 2010;70:283–289.

[13] Bessette P, Åslund F, Beckwith J, Georgiou G. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proceedings of the National Academy of Sciences 1999;96:13703–13708.

[14] Ballester M, Sánchez A, Folch J. Polymorphisms in the goat β-lactoglobulin gene. Journal of Dairy Research 2005;72:379–384.

[15] Alexander LJ, Hayes G, Pearse MJ. Complete sequence of the bovine β-lactoglobulin cDNA. Nucleic Acids Research 1989;17:6379.

[16] Loch J, Molenda M, Kopeć M, Świątek S, Lewiński K. Structure of two crystal forms of sheep β-lactoglobulin with EF-loop in closed conformation. Biopolymers 2014;101:886–894.

[17] Kontopidis G, Gilliver A, Sawyer L. Ovine β-lactoglobulin at atomic resolution. Acta Crystallographica, Section F: Structural Biology Communications 2014;70:1498–1503.

[18] Crowther JM, Lassé M, Suzuki H, Kessans SA, Loo TS, Norris GE, et al. Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin. FEBS Letters 2014;588:3816–3822.

[19] Loch J, Bonarek P, Polit A, Świątek S, Czub M, Ludwikowska M, et al. Conformational variability of goat β-lactoglobulin: crystallographic and thermodynamic studies. International Journal of Biological Macromolecules 2015;72:1283–1291.

[20] Oksanen E, Jaakola VP, Tolonen T, Valkonen K, Åkerström B, Kalkkinen N, et al. Reindeer β-lactoglobulin crystal structure with pseudo-body-centred noncrystallographic symmetry. Acta Crystallographica, Section D: Biological Crystallography 2006;62:1369–1374.

[21] Hoedemaeker FJ, Visschers RW, Alting AC, de Kruif KG, Kuil ME, Abrahams JP. A novel pH-dependent dimerization motif in β-lactoglobulin from pig (Sus scrofa). Acta Crystallographica, Section D: Biological Crystallography 2002;58:480–486.

[22] Brownlow S, Cabral J, Cooper R, Flower D, Yewdall S, Polikarpov I, et al. Bovine β-lactoglobulin at 1.8 Å resolution—still an enigmatic lipocalin. Structure 1997;5:481–495.

[23] Uhrínová S, Smith MH, Jameson GB, Uhrín D, Sawyer L, Barlow PN. Structural changes accompanying pH-induced dissociation of the β-lactoglobulin dimer. Biochemistry 2000;39:3565–3574.

[24] Kuwata K, Era S, Hoshino M, Forge V, Goto Y, Batt C. Solution structure and dynamics of bovine β-lactoglobulin A. Protein Science 1999;8:2541–2545.

[25] Fogolari F, Ragona L, Zetta L, Romagnoli S, Kruif KG, Molinari H. Monomeric bovine β-lactoglobulin adopts a β-barrel fold at pH 2. FEBS Letters 1998;436:149–154.

[26] Sakurai K, Goto Y. Dynamics and mechanism of the Tanford transition of bovine β-lactoglobulin studied using heteronuclear NMR spectroscopy. Journal of Molecular Biology 2005;356:483–496.

[27] Qin BY, Bewley MC, Creamer LK, Baker HM, Baker EN, Jameson GB. Structural basis of the Tanford transition of bovine β-lactoglobulin. Biochemistry 1998;37:14014–14023.

[28] Sakurai K, Goto Y. Manipulating monomer-dimer equilibrium of bovine β-lactoglobulin by amino acid substitution. Journal of Biological Chemistry 2002;277:25735–25740.

[29] McKenzie HA, Sawyer WH. On the dissociation of bovine β-lactoglobulins A, B, and C near pH 7. Australian Journal of Biological Sciences 1972;25:949–961.

[30] Bello M, Portillo-Téllez M, García-Hernández E. Energetics of ligand recognition and self-association of bovine β-lactoglobulin: differences between variants A and B. Biochemistry 2010;50:151–161.

[31] Mercadante D, Melton LD, Norris GE, Loo TS, Williams MAK, Dobson RCJ, et al. Bovine β-lactoglobulin is dimeric under imitative physiological conditions: dissociation equilibrium and rate constants over the pH range of 2.5–7.5. Biophysical Journal 2011;103:303–312.

[32] Stegen J, Ioannou J, Tromp H, Donald A, Schoot P. Mass-action driven conformational switching of proteins: investigation of β-lactoglobulin dimerisation by infrared spectroscopy. Journal of Physics D: Applied Physics 2015;48:384001.

[33] Kontopidis G, Holt C, Sawyer L. Invited review: β-lactoglobulin: binding properties, structure, and function. Journal of Dairy Science 2004;87:785–796.

[34] Ohtomo H, Konuma T, Utsunoiya H, Tsuge H, Ikeguchi M. Structure and stability of Gyuba, a β-lactoglobulin chimera. Protein Science 2011;20:1867–1875.

[35] Kobayashi T, Ikeguchi M, Sugai S. Construction and characterization of β-lactoglobulin chimeras. Proteins 2002;49:297–301.

[36] Pérez MD, Calvo M. Interaction of β-lactoglobulin with retinol and fatty acids and its role as a possible biological function for this protein: a review. Journal of Dairy Science 1995;78:978–988.

[37] Yang M, Guan H, Liu M, Lin Y, Yang J, Chen W, et al. Crystal structure of a secondary vitamin D3 binding site of milk β-lactoglobulin. Proteins 2008;71:1197–1210.

[38] Kontopidis G, Holt C, Sawyer L. The ligand-binding site of bovine β-lactoglobulin: evidence for a Function? Journal of Molecular Biology 2002;318:1043–1055.

[39] Wu SY, Pérez M, Puyol P, Sawyer L. β-Lactoglobulin binds palmitate within its central cavity. Journal of Biological Chemistry 1999;274:170–174.

[40] Loch J, Bonarek P, Polit A, Świątek S, Dziedzicka-Wasylewska M, Lewiński K. The differences in binding 12-carbon aliphatic ligands by bovine β-lactoglobulin isoform A and B studied by isothermal titration calorimetry and X-ray crystallography. Journal of Molecular Recognition 2013;26:357–367.

[41] Loch J, Bonarek P, Polit A, Riès D, Dziedzicka-Wasylewska M, Lewiński K. Binding of 18-carbon unsaturated fatty acids to bovine β-lactoglobulin—structural and thermodynamic studies. International Journal of Biological Macromolecules 2013;57:226–231.

[42] Loch J, Polit A, Bonarek P, Olszewska D, Kurpiewska K, Dziedzicka-Wasylewska M, et al. Structural and thermodynamic studies of binding saturated fatty acids to bovine β-lactoglobulin. International Journal of Biological Macromolecules 2012;50:1095–1102.

[43] Loch J, Polit A, Górecki A, Bonarek P, Kurpiewska K, Dziedzicka-Wasylewska M, et al. Two modes of fatty acid binding to bovine β-lactoglobulin—crystallographic and spectroscopic studies. Journal of Molecular Recognition 2011;24:341–349.

[44] Gutiérrez-Magdaleno G, Bello M, Portillo-Téllez CM, Rodríguez-Romero A, García-Hernández E. Ligand binding and self-association cooperativity of β-lactoglobulin. Journal of Molecular Recognition 2013;26:67–75.

[45] Loch J, Bonarek P, Polit A, Jabłoński M, Czub M, Ye X, et al. β-Lactoglobulin interactions with local anaesthetic drugs—crystallographic and calorimetric studies. International Journal of Biological Macromolecules 2015;80:87–94.

[46] Sawyer L, Brownlow S, Polikarpov I, Wu SY. β-Lactoglobulin: structural studies, biological clues. International Dairy Journal 1998;8:65–72.

[47] Ragona L, Fogolari F, Catalano M, Ugolini R, Zetta L, Molinari H. EF loop conformational change triggers ligand binding in β-lactoglobulins. The Journal of Biological Chemistry 2003;278:38840–38846.

[48] Yi C, Wambo TO. Factors affecting the interactions between beta-lactoglobulin and fatty acids as revealed in molecular dynamics simulations. Physical Chemistry Chemical Physics 2015;17:23074–23080.

[49] Narayan M, Berliner LJ. Mapping fatty acid binding to β-lactoglobulin: ligand binding is restricted by modification of Cys 121. Protein Science 1998;7:150–157.

[50] Mensi A, Choiset Y, Rabesona H, Haertlé T, Borel P, Chobert J-M. Interactions of β-lactoglobulin variants A and B with vitamin A. Competitive binding of retinoids and carotenoids. Journal of Agricultural and Food Chemistry 2013;61:4114–4119.

[51] Pérez MD, Puyol P, Ena JM, Calvo M. Comparison of the ability to bind lipids of β-lactoglobulin and serum albumin of milk from ruminant and non-ruminant species. Journal of Dairy Research 1993;60:55–63.

[52] Pérez MD, Sanchez L, Aranda P, Ena JM, Oria R, Calvo M. Effect of beta-lactoglobulin on the activity of pregastric lipase. A possible role for this protein in ruminant milk. Biochimica et Biophysica Acta 1992;1123:151–155.

[53] Seppälä M, Taylor R, Koistinen H, Koistinen R, Milgrom E. Glycodelin: a major lipocalin protein of the reproductive axis with diverse actions in cell recognition and differentiation. Endocrine Reviews 2002;23:401–430.

[54] Maier I, Okun V, Pittner F, Lindner W. Changes in peptic digestibility of bovine β-lactoglobulin as a result of food processing studied by capillary electrophoresis and immunochemical methods. Journal of Chromatography B 2006;841:160–167.

[55] Chaneton L, Sáez J, Bussmann L. Antimicrobial activity of bovine β-lactoglobulin against mastitis-causing bacteria. Journal of Dairy Science 2011;94:138–145.

[56] Hernández-Ledesma B, Recio I, Amigo L. β-Lactoglobulin as source of bioactive peptides. Amino Acids 2008;35:257–265.

[57] Pellegrini A, Dettling C, Thomas U, Hunziker P. Isolation and characterization of four bactericidal domains in the bovine beta-lactoglobulin. Biochimica et Biophysica Acta 2001;1526:131–140.

[58] Pan Y, Shiell B, Wan J, Coventry M. The molecular characterisation and antimicrobial properties of amidated bovine β-lactoglobulin. International Dairy Journal 2007;17:1450–1459.

[59] Oldfield DJ, Singh H, Taylor MW, Pearce KN. Heat-induced interactions of β-lactoglobulin and α-lactalbumin with the casein micelle in pH-adjusted skim milk. International Dairy Journal 2000;10:509–518.

[60] Henry G, Mollé D, Morgan F, Fauquant J, Bouhallab S. Heat-induced covalent complex between casein micelles and beta-lactoglobulin from goat's milk: identification of an involved disulfide bond. Journal of Agricultural and Food Chemistry 2002;50:185–191.

[61] Niemi M, Jylhä S, Laukkanen M-L, Söderlund H, Mäkinen-Kiljunen S, Kallio J, et al. Molecular interactions between a recombinant IgE antibody and the β-lactoglobulin allergen. Structure 2007;15:1413–1421.

Crowther Jennifer M.,Lassé Moritz,Suzuki Hironori,Kessans Sarah A.,Loo Trevor S.,Norris Gillian E.,Hodgkinson Alison J.,Jameson Geoffrey B. *and* Dobson Renwick C.J.*(*2014*),* **Ultra-high resolution crystal structure of recombinant caprine β-lactoglobulin***,* FEBS Letters*,* 588*,*
doi: http://dx.doi.org/10.1016/j.febslet.2014.09.010